



Revista Digital de Derecho Administrativo  
ISSN: 2145-2946  
Universidad Externado de Colombia

CAPDEFERRO VILLAGRASA, ÓSCAR; PONCE SOLÉ, JULI  
***Nudging e inteligencia artificial contra la corrupción en el sector público: posibilidades y riesgos***<sup>1</sup>

Revista Digital de Derecho Administrativo, núm. 28, 2022, pp. 225-258  
Universidad Externado de Colombia

DOI: <https://doi.org/10.18601/21452946.n28.08>

Disponible en: <https://www.redalyc.org/articulo.oa?id=503873276008>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

# *Nudging e inteligencia artificial contra la corrupción en el sector público: posibilidades y riesgos<sup>1</sup>*

ÓSCAR CAPDEFERRO VILLAGRASA<sup>2</sup>  
JULI PONCE SOLÉ<sup>3</sup>

## RESUMEN

El sector público está empezando a incorporar herramientas basadas en inteligencia artificial, entre las que debemos destacar las destinadas a la lucha contra la corrupción y el fraude como una de las actividades de las

- 1 El presente trabajo deriva de la ponencia titulada: "Acicates (*nudges*) digitales y regulación administrativa: inteligencia artificial y derecho a una buena administración en la gestión postpandemia", presentada por los dos autores en el XXVI Congreso Colombia 2021 del Centro Latinoamericano de Administración para el Desarrollo (CLAD). Asimismo, la investigación se enmarca en las actividades del proyecto de investigación PID2020-115774RB-I00, "Personalización de los servicios públicos, sesgos e inteligencia artificial: Hacia la consolidación de los derechos digitales en las administraciones públicas", IPs: Juli Ponce y Agustí Cerrillo, financiado por el Ministerio de Ciencia e Innovación español.
- 2 Doctor en Derecho por la Universidad de Barcelona, Barcelona, España. Profesor de Derecho Administrativo de la Universidad de Barcelona, Barcelona, España. Correo-e: ocapdeferro@ub.edu. Orcid: <https://orcid.org/0000-0001-9269-317>.
- 3 Doctor en Derecho por la Universidad de Barcelona, Barcelona, España. Profesor de Derecho Administrativo, Universidad de Barcelona, Barcelona, España. Correo-e: jponce@ub.edu. Orcid: <https://orcid.org/0000-0002-1977-5063>. Fecha de recepción: 21 de febrero de 2022. Fecha de modificación: 10 de abril de 2022. Fecha de aceptación: 5 de mayo de 2022. Para citar el artículo: CAPDEFERRO VILLAGRASA, ÓSCAR, y PONCE SOLÉ, JULI, "Nudging e inteligencia artificial contra la corrupción en el sector público: posibilidades y riesgos", *Revista digital de Derecho Administrativo*, Universidad Externado de Colombia, n.º 28, 2022, pp. 225-258. DOI: <https://doi.org/10.18601/21452946.n28.08>

administraciones más proclives al uso de herramientas basadas en inteligencia artificial. En el presente estudio se analiza cómo la unión de ciencias del comportamiento e inteligencia artificial pueden ser muy positivas para un adecuado control preventivo de la corrupción en la actuación de los servidores públicos, pero también abre la puerta a determinados riesgos propios del empleo de ambas técnicas (ciencias del comportamiento e inteligencia artificial), que deben tenerse en cuenta para así poder corregirlos.

**Palabras clave:** ciencias del comportamiento, buena administración, inteligencia artificial, corrupción, control administrativo, riesgos, políticas públicas.

## Nudging and Artificial Intelligence Against Corruption in the Public Sector: Possibilities and Risks

### ABSTRACT

The public sector is beginning to incorporate tools based on artificial intelligence, among which we must highlight those aimed at the fight against corruption and fraud. This study analyzes how the combination of behavioral sciences and artificial intelligence can be very helpful for an adequate preventive corruption control of public servants, but it also opens the door to certain risks inherent to the use of both techniques, which must be considered in order to correct them.

**Keywords:** Behavioral Sciences, Good Administration, Artificial Intelligence, Corruption, Administrative Control, Risks, Public Policies.

### INTRODUCCIÓN

El empleo de la inteligencia artificial en el sector público es una realidad que empieza a consolidarse y sobre la que se han publicado ya un número significativo de estudios<sup>4</sup>. Además, se debe destacar que la lucha contra la corrupción y el fraude ha sido una de las actividades de las administraciones

4 Entre otros, AGUSTÍ CERRILLO y MIQUEL PEGUERA, *Retos jurídicos de la inteligencia artificial*, Navarra: Aranzadi - Cizur Menor, 2020; y LORENZO COTINO, "Ética en el diseño para el desarrollo de una inteligencia artificial, robótica", *Revista catalana de dret públic*, n.º 58, 2019, pp. 29-48.

más proclives al uso de herramientas basadas en inteligencia artificial<sup>5</sup>: por citar solo dos ejemplos de interés, podemos mencionar los programas Arachne<sup>6</sup>, de la Comisión Europea, y Saler<sup>7</sup>, gestionado por la inspección de servicios valenciana. Asimismo, su desarrollo y uso parece que seguirá en aumento, puesto que se pueden encontrar recientes medidas para incentivar su uso, como se puede ver, en el caso español, en la Orden HFP/1030/2021, del 29 de septiembre, por la que se configura el sistema de gestión del Plan de Recuperación, Transformación y Resiliencia, en la que se establecen determinadas condiciones y requisitos para gestionar fondos europeos de recuperación pos-COVID-19<sup>[8]</sup>.

5 Véase ÓSCAR CAPDEFERRO, "Las herramientas inteligentes anticorrupción: Entre la aventura tecnológica y el orden jurídico", *Revista General de Derecho Administrativo*, n.º 50, 2019, pp. 1-28; JAVIER MIRANZO DÍAZ, "Inteligencia artificial y contratación pública", en Ignacio Martín y José Antonio Moreno Molina (dir.), *Administración electrónica, transparencia y contratación pública*, Madrid: Iustel, 2020, pp. 105-142.

6 El programa Arachne, desarrollado por la Comisión Europea, hasta la fecha se ofrece gratuitamente a las autoridades nacionales de los Estados miembros que gestionan fondos europeos y que manifiestan su interés en el programa. Este programa se desarrolló para dar cumplimiento a lo establecido por el Reglamento (UE) n.º 1303/2013 del Parlamento Europeo y del Consejo del 17 de diciembre de 2013, donde se establece la obligación para las autoridades de gestión de los Estados Miembros de "aplicar medidas antifraude eficaces y proporcionadas, teniendo en cuenta los riesgos detectados" (artículo 125.4.c del Reglamento). Este programa, basado en inteligencia artificial, analiza los riesgos de irregularidades en los expedientes de gestión de dichos fondos, entre los que se incluye la detección de posibles conflictos de intereses y la existencia de irregularidades previas de los sujetos participantes (como la condena por casos de corrupción), y facilita semanalmente un índice de riesgo que señalaría la posibilidad de que haya irregularidades en dichos expedientes. Al respecto, consultese <https://www.interact-eu.net/download/file/fid/6416> (última consulta: 14 de marzo de 2022).

7 Se trata de un sistema centrado en la detección de alertas, en el señalamiento temprano (esto es, mientras el correspondiente expediente administrativo todavía se encuentra abierto) de indicadores o indicios de irregularidad. Este sistema se aprobó mediante ley autonómica valenciana en 2018 (Ley 22/2018), y se puso en funcionamiento en febrero de 2019 para los expedientes de contratación pública, subvenciones y caja fija. Sobre este programa, véase JOSÉ DANIEL AMOEDO, "Los sistemas informáticos de detección de malas prácticas: Herramientas esenciales para la prevención de la corrupción", *Revista Internacional Transparencia e Integridad*, n.º 6, 2018; JULI PONCE, "La prevención de riesgos de mala administración y corrupción, la inteligencia artificial y el derecho a una buena administración", *Revista Internacional de Transparencia E Integridad*, n.º 6, 2018, pp. 1-19, y ALFONSO PUNCEL CHORNET, "Inteligencia artificial para la transparencia pública: El Sistema de Alertas Tempranas (Saler) de la Generalitat Valenciana", *Boletín económico de ICE*, n.º 3116, 2019, pp. 41-61.

8 Así, en dicha orden se establece que: "Las herramientas de detección del fraude pueden incluir: A. El uso de bases de datos como la Base Nacional de Datos de Subvenciones (BNDS), herramientas de prospección de datos ('data mining') o de puntuación de riesgos (ARACHNE)".

En el debate sobre el uso de este tipo de herramientas anticorrupción y antifraude, sin embargo, suelen dejarse de lado las posibles contribuciones que las ciencias del comportamiento pueden ofrecer para una mejor y más efectiva prevención de conductas contrarias a la ética pública y al marco normativo, lo cual no deja de ser hasta cierto punto sorprendente, dado el creciente interés que están generando este tipo de aportaciones en el diseño de políticas públicas<sup>9</sup>.

Así, algunas de las iniciativas más interesantes recientemente en materia de lucha contra la corrupción en el interior de las administraciones derivarían de la unión de las aportaciones de las ciencias de la conducta y los sistemas o programas de cumplimiento normativo (*compliance programs*), lo que ha dado lugar al enfoque sistemático de prevención de la corrupción y de promoción del comportamiento ético conocido como *behavioural public compliance*, y que puede trasladarse fácilmente y con gran efectividad a los entornos digitales en los que se empieza a trasladar el control de la corrupción, como señalábamos al inicio.

Sin embargo, no todos los efectos derivados del empleo de las ciencias de la conducta en los entornos digitales son positivos, ya que se identifican en la literatura científica algunos riesgos como la manipulación, casi inadvertida, de los individuos<sup>10</sup>.

Partiendo, pues, de esas constataciones iniciales, el objeto del presente estudio es delimitar en primer lugar cómo pueden las ciencias del comportamiento mejorar los sistemas (especialmente digitales) anticorrupción y, en segundo lugar, presentar qué riesgos se abren, a ser regulados o corregidos por los poderes públicos, al incorporar justamente las ciencias conductuales en los entornos digitales.

Entre las contribuciones de las ciencias del comportamiento, en este estudio nos queremos centrar particularmente en los *nudges* o acicates, como una de las principales derivadas de esas ciencias para la orientación de políticas públicas. Para entender el papel e importancia de esos *nudges* digitales, su potencialidad en la lucha contra la corrupción y su posible *lado oscuro*, hemos primero de entender el concepto y su contexto para luego aplicarlo al mundo digital y a las herramientas digitales basadas en inteligencia artificial para la lucha contra la corrupción y el fraude.

9 Por todos, véase BORJA BARRAGUÉ y JOSÉ NOGUERA, "Los nudges y el diseño conductual de políticas públicas", *Gestión y Análisis de Política Públicas*, n.º 25, 2021.

10 ISAAC BEN-ISRAEL, JORGE CERDIO et al., *Towards regulation of AI systems: Global perspectives on the development of a legal framework on Artificial Intelligence (AI) systems based on the Council of Europe's standards on human rights, democracy and the rule of law*, Consejo de Europa, 2020.

## 1. LAS APORTACIONES CONDUCTUALES Y LOS NUDGES

De acuerdo con las aportaciones provenientes de las denominadas *behavioural sciences* o ciencias de la conducta humana, sobre todo de la psicología, materia en la que destacan los conocidos trabajos de Kahneman, Premio Nobel de Economía en 2002<sup>[11]</sup>, se ha llegado a determinar que la racionalidad absoluta de la persona, del *homo economicus*, no existe, pues tal racionalidad está interferida por heurísticos y sesgos<sup>[12]</sup>. En definitiva, de estos avances se puede extraer, como afirma Pablo Grande, que aspectos como la presentación o el diseño de opciones presentadas a las personas van a influir orientando, en mayor o menor medida, su comportamiento<sup>[13]</sup>. Estos avances científicos son relevantes para las ciencias sociales en general y para el derecho en particular, si rechazamos para este un formalismo esterilizante y si, como indica, por ejemplo, el jurista italiano Casseste, en un artículo sobre los nuevos caminos para el derecho público<sup>[14]</sup>: "el derecho administrativo debe restablecer su lugar en el campo de las ciencias sociales". Ese lugar solo podrá ser ocupado con un planteamiento metodológico plural (legal, económico y político), si el derecho administrativo no quiere verse reducido a ser solo la tecnología de la litigiosidad judicial y pasar a ser irrelevante en fases previas a la judicial para el buen gobierno y la buena administración.

En la actualidad se habla ya de *economía y derecho conductuales*, puesto que las ciencias de la conducta humana llevan años, como decimos, cuestionando el modelo de *homo economicus* de la economía clásica, que ha influido también en el derecho público, basado históricamente en la regulación que piensa en el *hombre malo*, al que se refirió hace ya más de un siglo el jurista norteamericano Holmes en una célebre conferencia en la Universidad de Harvard, con racionalidad absoluta y maximizador de sus propios intereses, y como consecuencia, en el *orden y mando (command and control)*, esto es, en la prohibición y sanción como principales vías de intervención administrativa para garantizar el interés general y el cumplimiento de las normas.

Es a partir de las aportaciones científicas citadas cuando surge la noción de *nudging*, la cual ha inspirado comportamientos de las empresas privadas y concretas políticas públicas regulatorias en distintos países de todo el mundo,

11 DANIEL KAHNEMAN, *Pensar rápido, pensar despacio*, Barcelona: Debolsillo, 2013.

12 Se han identificado y clasificado un número significativo de sesgos. Por todos, véase OCDE, *Tools and Ethics for Applied Behavioural Insights: The Basic Toolkit*. OECD Publishing, 2019.

13 PABLO GRANDE SERRANO, "La psicología económica como herramienta para incentivar el cumplimiento voluntario de las obligaciones tributarias", *Revista de Contabilidad y Tributación*, n.º 439, 2019, pp. 18-19.

14 SABINO CASSESE, "New Paths for Administrative Law: A Manifesto", *International Journal of Constitutional Law*, vol. 10, n.º 3, 2012, pp. 603 y ss.

a las que el expresidente norteamericano Obama dedicó atención durante su mandato con diversas órdenes ejecutivas que dictó sobre este tema<sup>15</sup> y a las que bajo la presidencia de Biden también se ha aludido recientemente en un documento oficial<sup>16</sup>. Estas han dado lugar, además, a estructuras organizativas específicas para impulsarlas (caso notorio de la unidad británica Behavioural Team) e informes de la Unión Europea<sup>17</sup>.

Los *nudges* (concepto, este de *nudge*, que podría traducirse como pequeño empujón o codazo), en efecto, se han empleado en las políticas públicas a fin de incentivar o desincentivar determinadas conductas de la ciudadanía<sup>18</sup>. En su seminal trabajo sobre *nudging*, Richard Thaler (Premio Nobel de Economía en 2017) y Cass Sunstein (jurista norteamericano y gestor en temas de mejora regulatoria en la administración de Obama)<sup>19</sup> definen el *nudging* en el ámbito de las políticas públicas como<sup>20</sup> "cualquier aspecto de la arquitectura de las decisiones que modifica la conducta de las personas de una manera predecible sin prohibir ninguna opción ni cambiar de forma significativa sus incentivos económicos. Para que se pueda considerar *nudge*, debe ser barato y fácil de evitar. Los *nudges* no son órdenes".

Así, si nos centramos ahora en la vertiente pública administrativa del *nudging*, podemos afirmar que se trata de una técnica sofisticada, informada científicamente (a partir de las ciencias conductuales antes referidas), de

- 15 En ese sentido, véase THE WHITE HOUSE, Executive Order 13707 of September 15, Washington: Office of the Press Secretary, 2015 (Using Behavioral Science Insights to Better Serve the American People).
- 16 THE WHITE HOUSE, *Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking*, Washington: The White House. El 27 de enero, 2021, en este memorando se llama al uso de "Relevant approaches might include use of pilot projects, randomized control trials, quantitative-survey research and statistical analysis, qualitative research, ethnography, research based on data linkages in which records from two or more datasets that refer to the same entity are joined, well-established processes for community engagement and inclusion in research, and other approaches that may be informed by the social and behavioral sciences and data science".
- 17 UNIÓN EUROPEA, *Behavioural Insights Applied to Policy*, Luxemburgo: Publications Office of the European Union, 2016.
- 18 Así, Elena Costas y Pablo Tucat señalan que: "Los *nudges* en las políticas públicas implican usar el conocimiento derivado de las ciencias económicas, la neurociencia y la psicología, para incentivar o desincentivar actuaciones concretas de los distintos agentes, y alcanzar así objetivos establecidos por los dirigentes públicos. Al modificar el entorno en el que los ciudadanos toman decisiones, teniendo en cuenta los mecanismos de comportamiento que se activan en cada caso, los *nudges* constituyen un importante instrumento de políticas públicas en la acción de gobierno". ELENA COSTAS y PABLO TUCAT, "Nudges: diseño y evaluación", GAPP, n.º 25, 2021, p. 10.
- 19 Esta idea se encuentra basada en el proyecto de investigación PID2020-115774RB-I00 "Personalización de los servicios públicos, sesgos e inteligencia artificial: hacia la consolidación de los derechos digitales en las administraciones públicas", IPs: Juli Ponce y Agustí Cerrillo, financiado por el Ministerio de Ciencia e Innovación español.
- 20 En CASS SUNSTEIN y RICHARD TAHLER, *Un pequeño empujón (nudge)*, Madrid: Taurus, 2009.

fomento, a través de la cual gobiernos y administraciones buscan incentivar determinadas conductas alineadas con los intereses generales, pero sin emplear incentivos económicos. En empleo de esta técnica, los incentivos, que no serían económicos, se basarían en el conocimiento de heurísticos y sesgos. De forma orientativa, como indicación de tipologías, siguiendo a Cabrales y Rey, se puede afirmar que se pueden distinguir "cuatro clases de intervenciones tipo *nudge*: (1) provisión de información, (2) cambios en el entorno físico, (3) cambios en las opciones ofrecidas "por defecto" y (4) el uso de las normas sociales y de la relevancia de las opciones ofrecidas (*salience*)"<sup>21-22</sup>.

Este tipo de intervención administrativa, denominada a veces "paternalismo libertario" –en una expresión algo desafortunada, que bien podría substituirse por "intervención liberal", como ha sido sugerido<sup>23</sup>– aparece como interesante en el futuro práctico en nuestro propio contexto, para impulsar el derecho a una buena administración de los ciudadanos, mediante:

- a. La mejora de la calidad de las normas administrativas, en línea con el movimiento internacional de la *better regulation* (a través de análisis específicos incorporados en las evaluaciones de impacto normativo y mediante técnicas de mejora regulatoria).
- b. El uso inteligente de la publicidad activa (*smart disclosure*) realizada por las administraciones en el marco del cumplimiento de las leyes de transparencia. La publicidad activa pasa así de ser un simple, pero relevante, elemento de transparencia a ser una política pública en sí misma, que busca orientar la actividad del sector privado para la realización de actividades de interés general.
- c. La prevención de la corrupción de la mala administración negligente y la promoción del derecho a una buena administración (artículo 41 de la Carta de Derechos Fundamentales de la Unión Europea y normativa y jurisprudencia de la Unión Europea, de sus Estados miembros y de otros países<sup>24</sup>), mediante el diseño de adecuadas instituciones administrativas (organización y procedimiento).

21 ANTONIO CABRALES y PEDRO REY BIÉL, "Mas allá de los *nudges*: Políticas públicas efectivas basadas en la evidencia de las ciencias del comportamiento", *GAPP*, n.º 25, 2021, p. 40.

22 Esta idea se encuentra basada en el proyecto de investigación PID2020-115774RB-I00 ya mencionado.

23 SABINO CASSESE, "Exploring the legitimacy of Nudging", en Alexandra Kemmerer, Christoph Möllers, Maximilian Steinbeis y Gerhard Wagner (eds.), *Choice Architecture in Democracies. Exploring the Legitimacy of Nudging*, Baden-Baden: Nomos, 2016, pp. 241-246.

24 JULI PONCE, "La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida", *Cuadernos Democracia y Derechos Humanos*, n.º 15, 2019.

d. La personalización de los servicios públicos, que puede ser definida como la adaptación de la oferta a la demanda de un usuario a partir de los datos recogidos sobre sus hábitos. En definitiva, el objetivo prioritario sería prestar un mejor servicio y ofrecer un producto de acuerdo con los gustos o necesidades del usuario, para conseguir la individualización del producto y la personalización del servicio, con el fin de lograr la satisfacción del cliente e intentar orientar y ajustar, tanto como sea posible, el producto o el servicio a las necesidades de este<sup>25</sup>. Más en concreto, la personalización de los servicios públicos es una de las manifestaciones de la buena gestión pública y contribuye a hacer realidad el derecho a una buena administración de las personas. La personalización de servicios públicos se refiere a todo tipo de actuación administrativa, en el sentido, por ejemplo, dado en España a esta expresión, recogida con este significado en el artículo 106.2 de la Constitución, cuando se refiere al funcionamiento de los "servicios públicos"<sup>26</sup>.

En el ámbito del sector público se inscribe en lo que internacionalmente se conoce como el enfoque Citizen-Centric eGovernment<sup>27</sup>. A través de la personalización de los servicios públicos, estos se adaptan a las necesidades de las personas. Las administraciones públicas pueden así prestar los servicios públicos de manera proactiva anticipándose, de este modo, a la solicitud que pueda realizar el usuario o persona interesada, o sugiriéndole trámites que puedan ser de su interés.

La personalización y la anticipación en la prestación de los servicios públicos están estrechamente relacionadas y constituyen dos de los principales mecanismos identificados para la innovación de los servicios públicos. Ambos procesos permiten simplificar los procedimientos de acreditación del cumplimiento de los requisitos para disfrutar de un mejor servicio público.

e. Una intervención pública eficaz y barata en la conducta de los privados en el mercado, en garantía de los intereses generales, aspecto que ha centrado en buena medida los análisis existentes en la materia.

25 AMPARO FRÍAS, "Los servicios personalizados de información de actualidad de los medios de comunicación españoles a través de Internet". *BiD*, n.º 17, 2006.

26 "Los particulares, en los términos establecidos por la ley, tendrán derecho a ser indemnizados por toda lesión que sufran en cualquiera de sus bienes y derechos, salvo en los casos de fuerza mayor, siempre que la lesión sea consecuencia del funcionamiento de los servicios públicos".

27 COMISIÓN EUROPEA, *A Handbook for Citizen-centric eGovernment*, Ecotec, 2007.

## 2. LOS NUDGES DIGITALES

Pues bien, en relación con todo lo dicho, la tecnología puede permitir incrementar la diligencia debida administrativa exigida por el derecho a una buena administración al personalizar la prestación de servicios públicos, permitiendo su adaptación y actuación proactiva en favor de la ciudadanía, así como también prevenir la corrupción y la mala administración, como veremos más adelante.

En cuanto a la personalización antes aludida, las administraciones públicas han encontrado históricamente barreras para poder personalizar la prestación de los servicios públicos por diversos motivos como el gran volumen de destinatarios y usuarios, la falta de recursos (humanos, económicos, tecnológicos) e incluso en ocasiones el propio ordenamiento jurídico (por ejemplo, en materia de protección de datos personales o de seguridad de la información). Sin embargo, en la actualidad, el desarrollo de la tecnología puede facilitar el proceso de la personalización y la anticipación en la prestación de los servicios públicos y reducir significativamente algunas de las barreras que tradicionalmente han dificultado este proceso no siendo la masificación ya un obstáculo.

Ahora bien, aunque no cabe duda de que la tecnología puede ayudar a hacer eficaz el derecho a una buena administración de los ciudadanos, mediante esa personalización aludida y mediante la prevención de la mala administración y de la corrupción administrativa, tampoco hay dudas en cuanto al peligro que, a la vez, y paradójicamente, puede presentar para la buena administración<sup>28</sup>.

Así, en el ámbito de la inteligencia artificial (IA), los algoritmos tienen el potencial de promover la eficacia, la eficiencia, la consistencia y la justicia, pero también pueden reforzar discriminaciones históricas u obscurecer comportamientos indeseables<sup>29</sup>. Existen datos de que clasificadores algorítmicos de riesgo usados en algunos estados de los EE. UU. en el ámbito de la justicia doblan la clasificación de personas negras como futuros criminales en relación con las personas blancas de forma incorrecta.

Pero los sesgos algorítmicos no son el único problema. Tanto en el entorno digital como fuera de él se llevan a cabo continuamente arquitecturas de

28 Para más detalle, JULI PONCE, "La lucha por el buen gobierno y el derecho a una buena administración", *op. cit.*

29 Véase la Algorithmic Justice League en <https://www.ajlunited.org/>. También House of Commons Science and Technology Committee, *Algorithms in decisionmaking*, Fourth Report of Session 2017-2019, 2018. Disponible en: <https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/351/35102.htm>.

la elección<sup>30</sup>, a las que se han referido Thaler y Sunstein<sup>31</sup>, sea por activa o por pasiva. Recordemos que se habla de arquitectura de la elección porque cuando se le pide a alguien que haga una elección, la forma en que se ofrecen las alternativas afecta a su respuesta. Los seres humanos se enfrentan a elecciones todos los días, pero el resultado de cualquier elección está influido no solo por las deliberaciones racionales de las opciones disponibles, sino también por el diseño del entorno de elección en el que se presenta la información, que puede ejercer una influencia subconsciente en el resultado. En otras palabras, lo que se elige a menudo depende de cómo se presenta la elección, de modo que la arquitectura de la elección altera el comportamiento de las personas de forma predecible. Incluso simples modificaciones del entorno de elección en el que se presentan las opciones pueden influir en las elecciones de las personas y “empujarlas” a comportarse de determinadas maneras. De hecho, no existe una forma neutral de presentar las opciones. Por ejemplo, se ha podido constatar que el empleo de técnicas de *nudging*, en particular cambiando las opciones por defecto (de *opt-in* a *opt-out*) en el contexto de la donación de órganos, ha resultado en un incremento considerable del número de donantes<sup>32-33</sup>.

El *nudging* digital es, entonces, el uso de elementos de diseño de interfaz de usuario para guiar el comportamiento de las personas en entornos de elección digital. Los entornos de elección digital son interfaces de usuario –como formularios web y pantallas de ERP (*enterprise resource planning*)– que requieren que las personas hagan juicios o tomen decisiones.

En concreto, en el ámbito administrativo, existe, pues, el riesgo de que el enfoque *citizen-centric*, al que nos hemos referido, derive, en realidad, hacia una manipulación, en la que se utilice por las Administraciones los datos masivos y la IA no ya simplemente para personalizar servicios, sino para manipular a las personas en relación con estos usando *nudges* digitales<sup>34</sup>.

30 RICHARD THALER, CASS SUNSTEIN y JOHN BALZ, “Choice Architecture”, *The Behavioral Foundations of Public Policy*, n.º 25, 2012.

31 CASS SUNSTEIN y RICHARD TAHLER, *Un pequeño empujón*, op. cit.

32 Así se recoge, entre otros autores, por Costas y Tucat, con cita de varios estudios. ELENA COSTAS y PABLO TUCAT, “Nudges: diseño y evaluación”, op. cit., pp. 14 y 18.

33 Esta idea se encuentra basada en el proyecto de investigación PID2020-115774RB-I00 ya mencionado.

34 SOFIA RANCHORDÁS, “Nudging citizens through technologies in Smart Cities”, *International Review of Law, Technology and Computers*, n.º 33, 2019, expone diversos ejemplos de *nudging smart cities*. Así, por ejemplo, Durham, en Carolina del Norte, en colaboración con empresas, ha empleado *nudges* basados en datos recopilados para incentivar la elección por parte de los ciudadanos del transporte público en vez de los vehículos privados, mediante el envío de correos electrónicos personalizados con rutas de los domicilios al trabajo, mostrando diferentes opciones de transporte público y privado. Boston ha introducido una aplicación para el *smartphone* (Boston’s Safest Driver) que proporciona

Pero también existe el riesgo en el ámbito del sector privado de que la actividad económica conduzca a una manipulación de los ciudadanos<sup>35</sup>, con la obtención de datos personales mediante trickeyuelas y engaños y la explotación de sus sesgos cognitivos para forzar un mayor consumo mediante el *nudging* digital.

El Comité de Ministros del Consejo de Europa, en una declaración del 13 de febrero de 2019, ha alertado sobre el riesgo aquí apuntado<sup>36</sup>.

Sin duda las posibilidades ofrecidas por la inteligencia artificial (IA) hacen aún mayor el riesgo de manipulación (y de discriminación) mediante el

*feedback* sobre la conducción basado en la velocidad, aceleración, frenadas y otros elementos derivados de la recolección de datos urbanos y personales de la seguridad vial y recompensa a los conductores cuidadosos con premios semanales al conductor más seguro. En Eindhoven, una calle (*Stratumseind*, calle de la vida nocturna) está llena de localizadores de wifi, cámaras y micrófonos que pueden detectar comportamientos agresivos, y cuenta con una iluminación urbana inteligente que cambia la intensidad de la luz para calmar a las personas. La ciudad también prevé difundir aromas relajantes para alterar el estado de ánimo de la gente que pudiera potencialmente iniciar altercados nocturnos.

35 SARAH ZIMMERMANN, "¿Incentivo o manipulación?", *Mente y Cerebro*, n.º 78, 2016, pp. 17 y ss.

36 Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes, adopted by the Committee of Ministers on 13 February 2019 at the 1337 meeting of the Ministers' Deputies. La traducción del inglés es nuestra: "La conciencia pública, sin embargo, sigue siendo limitada en cuanto a la medida en que los dispositivos cotidianos reúnen y generan grandes cantidades de datos. Esos datos se utilizan para capacitar a las tecnologías de aprendizaje automático (*machine-learning*) a fin de dar prioridad a los resultados de las búsquedas, predecir y configurar las preferencias personales, alterar los flujos de información y, a veces, someter a los individuos a experimentos de comportamiento. [...] Las herramientas contemporáneas de aprendizaje automático tienen la creciente capacidad no solo de predecir las elecciones sino también de influir en las emociones y los pensamientos, y alterar un curso de acción anticipado, a veces de manera subliminal. Los peligros para las sociedades democráticas que se derivan de la posibilidad de emplear esa capacidad para manipular y controlar no solo las opciones económicas sino también los comportamientos sociales y políticos, solo se han hecho evidentes recientemente. En este contexto, debe prestarse especial atención al importante poder que el avance tecnológico confiere a quienes –ya sean entidades públicas o agentes privados– pueden utilizar esos instrumentos algorítmicos sin una supervisión o un control democráticos adecuado [...]. considerando la necesidad de marcos de protección adicionales relacionados con los datos que vayan más allá de las nociones actuales de protección de los datos personales y la privacidad y que aborden los importantes efectos del uso selectivo de los datos en las sociedades y en el ejercicio de los derechos humanos en general; [...] Iniciar, dentro de los marcos institucionales apropiados, debates públicos abiertos, informados e inclusivos con miras a proporcionar orientación sobre dónde trazar la línea entre las formas de persuasión permisibles y la manipulación inaceptable. Esta última puede adoptar la forma de una influencia subliminal, explotar las vulnerabilidades o los sesgos cognitivos existentes o invadir la independencia y la autenticidad de la adopción de decisiones individuales".

denominado *nudge de precisión* o *hyper nudge*. Un reciente documento impulsado también por el Consejo de Europa se ha referido también a esta realidad<sup>37</sup>, señalando que: "También hay que prestar especial atención al uso potencial de la IA en la interacción hombre-máquina para aplicar estrategias de *nudging*. En este caso, debido a la complejidad y oscuridad de las soluciones técnicas adoptadas, la IA puede aumentar el papel pasivo de los ciudadanos y afectar negativamente al proceso democrático de toma de decisiones. Por el contrario, un enfoque activo basado en la participación consciente y la participación activa en los objetivos de la comunidad debería ser preferido y mejor gestionado por las herramientas de participación de la IA. Cuando se adopten, las estrategias de *nudging* deben seguir un enfoque basado en la evidencia" (la traducción del inglés es nuestra).

El uso de estas tecnologías, alguna de las cuales incluye inteligencia artificial, puede ser sin duda interesante para el logro del derecho a una buena administración, pero también presentan dificultades de aplicación, que pueden ser salvados mediante conocimientos conductuales y riesgos indudables para los derechos de las personas. En cuanto a los riesgos, estos pueden estar presentes en aquellos casos en que en el entorno digital se emplean *nudges* para incentivar determinados comportamientos de consumidores o usuarios de servicios públicos.

Analizaremos aquí, brevemente, como esos incentivos se están desarrollando en la práctica en el sector privado y en el público, mediante su uso en plataformas digitales. Trataremos, en primer lugar, cómo puede emplearse este tipo de incentivos en el marco de los programas de lucha contra la corrupción y la promoción de la buena administración (con especial atención a los programas de cumplimiento normativo conductual del sector público), y a continuación abordaremos algunos de los riesgos que ha identificado la literatura, con carácter general, en el uso de *nudges* en entornos digitales, los cuales, como se verá en el apartado 4.3, también pueden llegar a materializarse, hasta cierto punto, en el sector público, lo que podría dar lugar a un tipo de control sobre la actuación administrativa acaso excesivo, con pocas garantías y desproporcionado.

Empecemos, como decíamos, por estudiar las posibilidades en el empleo de acicates o *nudges* por parte de las administraciones para promover la actuación ética y prevenir la corrupción.

37 CONSEJO DE EUROPA, *Towards regulation of AI systems*, op. cit.

### 3. EL USO DE NUDGES O ACICATES POR PARTE DE LAS ADMINISTRACIONES PÚBLICAS PARA LA BUENA ADMINISTRACIÓN: LA PREVENCIÓN DE LA CORRUPCIÓN Y EL CUMPLIMIENTO NORMATIVO PÚBLICO CONDUCTUAL (BEHAVIOURAL PUBLIC COMPLIANCE)

Como ya avanzamos antes, el uso de *nudges* digitales, en su caso basados en la IA, puede conducir a una mejor prevención de la corrupción y de la mala administración. La combinación de programas de cumplimiento normativo en el sector público, acompañados de técnicas derivadas de las ciencias conductuales, como podría ser la incorporación de *nudges*, puede resultar muy efectiva en la lucha contra la corrupción y en la garantía de la ética pública.

Las ciencias conductuales han abordado, específicamente, las causas que conducen a comportamientos contrarios a la ética en las organizaciones<sup>38</sup>, como podría ser, por supuesto, una Administración pública. Entre estas conductas contrarias a la ética se cuenta como una de las más graves formas de quebrantamiento de las reglas de la ética pública la corrupción. Además, con base en algunos estudios empíricos centrados en el análisis de los distintos tipos de violaciones éticas que se denuncian e investigan internamente, llevados a cabo en distintas administraciones en países europeos, se puede señalar que, aunque la corrupción no es el tipo de infracción ética más frecuente<sup>39</sup>, igualmente ocuparía un lugar destacado, ya que los casos de corrupción se sitúan, en estos trabajos, en torno a una cuarta parte del total de las infracciones éticas investigadas<sup>40</sup>.

Podemos entender, con carácter general, la corrupción pública como el abuso de la posición o cargo público en beneficio privado<sup>41</sup>, lo que constituye, como ya hemos avanzado, una mala práctica<sup>42</sup> y una grave vulneración de los

- 38 Una revisión de la literatura de interés se encuentra en ERIC BRENDS *et al.*, *Unethical behaviour in the workplace: a rapid assessment of the research literature*, CIPD: Technical Report, 2019.
- 39 Serían más frecuentes el comportamiento inapropiado y el despilfarro, y en especial, la realización de tareas no relacionadas con el trabajo durante el horario laboral, la intimidación o amenaza a compañeros.
- 40 GJALT DE GRAAF, LEO HUBERTS Y TEBBINE STRÜWER, "Integrity violations and corruption in Western public governance: Empirical evidence and reflection from the Netherlands", *Public Integrity*, vol. 20, n.º 2, 2018, pp. 131-149.
- 41 OFICINA ANTIFRAUDE DE CATALUÑA, *Glosario de Integridad*, 2018. Disponible en: <https://www.antifrau.cat/es/la-corrupcion/glosario.html>.
- 42 Sobre este concepto, resulta de interés el preámbulo de la Ley 22/2018, del 6 de noviembre, de Inspección General de Servicios y del sistema de alertas para la prevención de malas prácticas en la Administración de la Generalitat y su sector público instrumental, donde se aporta la siguiente definición: "Las malas prácticas, entendidas como

valores constitucionales definitorios de la Administración (así, artículo 103 Constitución española), de la buena administración y del buen gobierno<sup>43</sup>.

En la actualidad, como enseguida se desarrollará, uno de los pilares en la lucha contra la corrupción es el enfoque de cumplimiento normativo o *compliance*, basado en la adopción de programas o sistemas de cumplimiento, y este enfoque presenta un espacio idóneo para la incorporación de los conocimientos provenientes de las ciencias conductuales, para promover en el ejercicio de las funciones públicas un comportamiento alineado con la ética pública y plenamente ajustado al marco jurídico, evitando prácticas ilícitas o incumplimientos éticos. Para ello, primero expondremos qué es el sistema de *compliance* y por qué es importante; a continuación, prestaremos atención a la introducción de las ciencias del comportamiento en los sistemas de *compliance*; y por último, nos centraremos en exponer cómo pueden encontrar, así, el *nudging* y el ya mencionado *hyper nudging*, un espacio en el servicio público, en prevención de la corrupción y otras malas prácticas afines.

### 3.1. EL CUMPLIMIENTO NORMATIVO Y LA PREVENCIÓN DE LA CORRUPCIÓN Y LA MALA ADMINISTRACIÓN

Con el objeto de lograr una adecuada prevención de la corrupción y otras malas prácticas administrativas, así como para garantizar la integridad pública, el servicio público debe mostrarse permeable a las nuevas técnicas y metodologías que, desde hace ya bastantes años, vienen impulsando y recomendando las principales instancias internacionales con competencias en la lucha contra la corrupción. Esas técnicas, en términos generales, se basan en la identificación, el análisis y la gestión de riesgos de corrupción (y de otras irregularidades), como piedra angular de un sistema verdaderamente garante de la integridad y efectivo contra la corrupción<sup>44</sup>. Así lo señala, por ejemplo, la Oficina de las Naciones Unidas contra la Droga y el Delito, perteneciente a la Organización de las Naciones Unidas, en una guía de 2020 elaborada para promover entre los Estados la adopción de un sistema de

aquellas actuaciones de la administración que puedan propiciar el incumplimiento de los principios de legalidad, objetividad, imparcialidad, eficacia y eficiencia, así como las posibles irregularidades administrativas, afectan no solo al prestigio y la legitimidad de cualquier organización, sino que también merman de manera sustancial la eficacia y eficiencia de sus actuaciones".

- 43 Conceptos, estos, vinculados también con la integridad en el servicio público. Así, por ejemplo, Ley 4/2011, artículo 3.º, literal K, de la buena administración y del buen gobierno de las Illes Balears.
- 44 OCDE, *Towards a Sound Integrity Framework: Instruments, Processes, Structures and Conditions for Implementation*, GOV/PGC/GF(2009)1, OCDE, 2009, p. 21.

gestión de riesgos de corrupción<sup>45</sup>; o la OCDE, que ha incidido en la relevancia de un enfoque sistémico basado en el análisis y gestión de riesgos para la integridad en la configuración de un buen sistema de integridad pública, en su recomendación sobre integridad pública de 2017<sup>[46]</sup>.

Es, pues, en este contexto de recomendaciones internacionales para adoptar sistemas anticorrupción y de integridad pública basados en la gestión de riesgos cuando surgen, como herramientas de interés para el sector público en su tarea de garantizar la integridad pública, los programas de cumplimiento o *compliance*<sup>47</sup>, en los que sobre la base de una identificación de riesgos previa, se establece todo un sistema o infraestructura potencialmente capaz de asegurar el comportamiento ajustado a derecho y a la ética.

Los programas de cumplimiento, principalmente, han tenido acogida y desarrollo en el sector privado, en el que han sido adoptados por las empresas, sobre todo con objeto de reducir o eludir responsabilidades penales<sup>48</sup>. Si bien el contenido de estos programas puede ser variable, se puede destacar que, entre los mismos, una pieza relevante es el código ético y de conducta, que, junto con el proceso de gestión de riesgos, constituiría el eje central de este tipo de programas<sup>49</sup>. Otros elementos que se consideraría necesario incluir serían los canales de denuncia y los organismos de control<sup>50</sup>.

En efecto, se ha considerado que, aun siendo la gestión de riesgos la pieza central para garantizar la integridad, reforzar el buen gobierno y prevenir la corrupción, son necesarios otros elementos para que el sistema preventivo funcione correctamente. Esta idea es la que ha inspirado los sistemas de cumplimiento (o *compliance*) y los sistemas de gestión anticorrupción estandarizados a través de las normas técnicas ISO 19600:2014<sup>[51]</sup> (*Compliance*

45 UNODC, *State of Integrity: A guide on conducting corruption risk assessments in public organizations*, Viena: Naciones Unidas, 2009.

46 OCDE, *Recomendación del Consejo de la OCDE sobre integridad pública*, OCDE, 2017, p.12.

47 CONCEPCIÓN CAMPOS ACUÑA, *Compliance en la Administración pública: Dificultades y propuestas*, Granada: CEMCI, 2017; y JOAN QUERALT, "Public compliance y corrupción: Análisis conceptual y propuestas", *Revista Internacional de Transparencia e Integridad*, n.º 2, 2016.

48 Por ejemplo, artículo 31 bis del Código Penal español (Ley Orgánica 10/1995, del 23 de noviembre).

49 ADÁN NIETO MARTÍN, "De la ética pública al *public compliance*: Sobre la prevención de la corrupción en las Administraciones públicas", en Adán Nieto Martín (dir.), *Public compliance. Prevención de la corrupción en administraciones públicas y partidos políticos*, Madrid: UCLM, 2014, pp. 17-42, 23.

50 CONCEPCIÓN CAMPOS ACUÑA, *Aplicación práctica del compliance en la contratación pública*, Pamplona: Aranzadi, 2019, pp.195 y ss.

51 ISO, *Compliance management systems. Guidelines*, 2021. Disponible en: <https://www.iso.org/standard/62342.html>.

*management systems. Guidelines)<sup>52</sup> y 37001:2016 (Anti-bribery management systems. Requirements with guidance for use).*

No forma parte del objeto ni de las intenciones de este trabajo exponer aquí los detalles metodológicos y estructurales de los sistemas de cumplimiento, cuestión sobre la que nos remitimos a la citada normativa técnica y a la abundante bibliografía sobre *compliance* existente<sup>53</sup>. Sin embargo, resulta útil resaltar que este tipo de sistemas o programas tiene por objeto principal detectar las vulnerabilidades de la organización (aquellos ámbitos o actuaciones donde existen mayores riesgos de incumplimiento normativo o de quebrantamiento de los principios y valores éticos de la organización) y establecer, en consecuencia, medidas de tratamiento de tales vulnerabilidades y de control del cumplimiento, para lo que se sirven de una serie de herramientas (mapas de riesgos, listas de comprobación, canales de denuncia internos para alertar de posibles incumplimientos, órganos de control y supervisión, etc.).

Presentado, así, el sistema de cumplimiento normativo, ya podemos identificar cómo encajan en éste las aportaciones de las ciencias conductuales.

### 3.2. EL CUMPLIMIENTO NORMATIVO Y LAS APORTACIONES CONDUCTUALES: EL *BEHAVIOURAL PUBLIC COMPLIANCE*

Los conocimientos generados en el campo de las ciencias conductuales, particularmente los referidos a la promoción de los comportamientos éticos, se pueden integrar en los sistemas de *compliance*, sin lugar a duda, especialmente en el momento de diseñar medidas efectivas para mitigar los riesgos o vulnerabilidades detectadas, y que favorezcan en la organización conductas ajustadas a las normas (jurídicas y éticas). De esta unión ha surgido, con cada vez mayor arraigo en la literatura científica, el concepto de *behavioral compliance*, el cual, como ya hemos avanzado, se puede entender como la adopción, en los sistemas de cumplimiento, de enfoques basados en la evidencia científica ofrecida por las ciencias conductuales<sup>54</sup>, especialmente

52 En ámbito penal y para el Estado español, resulta pertinente la certificable UNE 19601:2017, *Sistemas de gestión de compliance penal. Requisitos con orientación para su uso*. Disponible en: <https://www.aenor.com/normas-y-libros/buscadorm-de-normas/une?c=N0058338>.

53 Por todos, con especial atención al *compliance* aplicado al sector público, véase CONCEPCIÓN CAMPOS ACUÑA, *Guía práctica de compliance en el sector público*, Madrid: Walters Kluwer, 2020.

54 DANIEL TEJADA PLANAS, *Behavioral Compliance: reforzando el compliance a través de la ética conductual*, Cizur Menor: Thomson Reuters Aranzadi, 2021; y DONALD LANGEVOORT, "Behavioral Ethics, Behavioral Compliance", en JENNIFER ARLEN, *Research Handbook on Corporate Crime and Financial Misdealing*, Edward Elgar Publishing, 2018.

para el diseño de medidas efectivas de tratamiento de los riesgos o vulnerabilidades detectados<sup>55</sup>.

En la actualidad se está desarrollando investigación en un campo que ha sido visto con especial interés para reforzar los sistemas de cumplimiento normativo, conocido como *behavioral ethics* o ética conductual. La ética conductual ha sido definida como un campo emergente de la investigación científica dedicada al estudio de los aspectos cognitivos de la toma de decisiones éticas<sup>56</sup>, habitualmente centrado en el contexto de una organización como marco en el que las personas adoptan comportamientos ajustados o no a la ética<sup>57</sup>, y que asumiría, como punto de partida, que existen factores que facilitan que cualquier persona pueda acabar realizando acciones contrarias a las normas jurídicas y a la ética<sup>58</sup>.

Así, existen múltiples estudios de interés que abordan de forma empírica el impacto de determinadas medidas para promover el comportamiento ético en una organización<sup>59</sup>, de los que se pueden extraer relevantes orientaciones para el diseño de las medidas de tratamiento de los riesgos de comportamientos ilícitos o no éticos detectados en la organización.

Como dijimos en el apartado anterior, exponiendo el caso británico, existe un debate sobre la efectividad de la formación en materia de ética para evitar sesgos cognitivos. Pero, desde una perspectiva más general, por ejemplo, Van Montfort y otros<sup>60</sup> concluyeron, tras su estudio, que la formación ética tiene efectos en el grado de cumplimiento de los valores y principios éticos a corto plazo, pero no a largo plazo. Esta aportación, basada en la evidencia científica, empírica, sin duda, podría servir para orientar

55 DANIEL TEJADA PLANA, *Behavioral Compliance*, op. cit., pp. 84-86.

56 YUVAL FELDMAN y YOTAM KAPLAN, "Big Data and bounded ethicality", *Cornell Journal of Law and Public Policy*, vol. 29, 2019, p. 41.

57 LINDA TREVINO, GARY WEAVER y SCOTT REYNOLDS, "Behavioral Ethics in Organizations: A Review", *Journal of Management*, vol. 32, n.º 6, 2006, p. 952.

58 Así, TODD HAUGH, "Nudging Corporate Compliance", *American Business Law Journal*, vol. 54, n.º 4, 2017, p. 707) afirma que: "the central finding of the many studies conducted under the behavioral ethics umbrella—that 'cognitive heuristics, psychological tendencies, social and organizational pressures, and even seemingly irrelevant situational factors can make it more likely that good people will do bad things'. In human terms, behavioral ethics research tells us that while most people are moral individuals intent on doing right, we are not as ethical as we think we are. [...] Most people will make moral decisions in line with their ethical beliefs, but only to a point. Because of cognitive obstacles, which may be exacerbated by external factors, 'many people are blind to their own unethical conduct'".

59 Como JULIA THALER y BERND HELMIG, "Do codes of conduct and ethical leadership influence public employees' attitudes and behaviours? An experimental analysis", *Public Management Review*, vol. 18, n.º 9, 2016, pp. 1365-1399; o recopilados en revisiones de literatura previa, entre las que podemos destacar el estudio ya citado de ERIC BARENDS et al., *Unethical behaviour in the workplace*, op. cit.

60 ANDRÉ VAN MONTFORT, LAURA BECK y ANNEKE TWINIJSTRA, "Can integrity be taught in public organizations?", *Public Integrity*, vol. 15, n.º 2, 2013, pp. 117-132.

las medidas de cumplimiento normativo determinando que el programa de formación en ética y cumplimiento debe incorporarse en la organización como una actividad de realización periódica (dado su efecto beneficioso limitado solo al corto plazo), especialmente para aquellas personas en cargos y puestos particularmente expuestos a riesgos de malas prácticas. En una línea parecida, Hauser<sup>61</sup> identifica el beneficio que se puede obtener con los programas de formación para que los empleados rechacen las prácticas corruptas o presenten una menor tendencia a justificarlas, y destaca que para ser particularmente útiles deben prestar especial atención a las zonas grises o cuestiones acaso ambiguas, poco claras, en materia de ética y cumplimiento normativo, de donde se podría extraer que la formación debe ser concreta, práctica y centrada en resolver cuestiones que la normativa aplicable no resuelva de manera clara<sup>62</sup>.

Otras contribuciones de interés para el efectivo empleo de medidas que incentiven el cumplimiento y la ética derivan de la constatación, mediante experimentos, de que un recordatorio ético en el momento previo a la toma de decisión reduce el número de incumplimientos. Se ha podido comprobar que había menos comportamientos deshonestos cuando se pedía, justo antes de realizar el experimento, que los participantes recordaran los Diez Mandamientos, que firmaran un código ético e, incluso, que firmaran una cláusula de veracidad o cumplimiento justo antes de llenar un formulario (factor, este último, que se demostró mucho más efectivo que incluir esa cláusula al final, puesto que en este último caso se activaría el comportamiento moral en un momento en que ya no incidiría sobre la tarea)<sup>63</sup>. De ese conjunto de experimentos se podría extraer que es efectivo incluir mensajes de sensibilización o recordatorios éticos y de legalidad justo antes de la adopción de determinadas actuaciones, en las que concurren significativos riesgos de mala administración.

Esta última propuesta, además, puede vincularse con el *nudging*, como enseguida se verá.

### 3.3. EL CUMPLIMIENTO NORMATIVO CONDUCTUAL Y LOS NUDGES DIGITALES

El *nudging*, como ya se expuso, es una aproximación al diseño de la toma de decisión tendente a incentivar un determinado tipo de decisión (en este caso, la ajustada a la legalidad y a la ética pública), basándose en el conocimiento

61 CHRISTIAN HAUSER, "Fighting Against Corruption: Does Anti-corruption Training Make Any Difference?", *Journal of Business Ethics*, vol. 159, n.º 1, 2019, pp. 281-299.

62 Para más información sobre formación efectiva en materia de ética y cumplimiento, con revisión de la literatura y aplicando *behavioral ethics*, véase CHRISTIAN HAUSER, "From Preaching to Behavioral Change: Fostering Ethics and Compliance Learning in the Workplace", *Journal of Business Ethics*, vol. 162, n.º 4, 2020, pp. 835-855.

63 DANIEL TEJADA PLANA, *Behavioral compliance*, op. cit., pp. 102-104.

que las ciencias conductuales ofrecen sobre los aspectos que inciden en la toma de decisiones. En consecuencia, el *nudging* forma parte del anteriormente expuesto *behavioral compliance*, y se manifiesta en la adopción de medidas dirigidas a operar en el momento de tomar decisiones particularmente expuestas a riesgos de corrupción u otras formas de mala administración, promoviendo la adopción de conductas ajustadas a las normas jurídicas y éticas aplicables.

La OCDE<sup>64</sup> ha sugerido el empleo de distintas herramientas para modificar la arquitectura de la decisión de los empleados públicos en favor de la integridad. Así, sugiere el uso de opciones por defecto en los procedimientos (por ejemplo, al comienzo de un nuevo proyecto, cada miembro del equipo recibe automáticamente un formulario para declarar cualquier conflicto de intereses), el empleo de recordatorios y encuadres (incluyendo mensajes morales, por ejemplo, sobre regalos, o personalizados, que refuerzan la identidad principal, o que acudan a normas sociales), avanzar en la simplificación y comodidad (por ejemplo, presentación de la declaración de bienes a través del sistema electrónico y al mismo tiempo que las declaraciones de impuestos; uso de un botón rojo para presentar una decisión sensible), *gaming* (por ejemplo, organización de juegos, concursos, loterías, etc., con microincentivos, por ejemplo un desayuno en la oficina), incentivación de compromisos (por ejemplo, autocompromisos éticos: "Declaro que toda la información introducida a continuación será veraz") o entrenamiento en dilemas éticos.

En esta línea podemos pensar, por ejemplo, en el siguiente tipo de *nudges* o acicates, empleados en el marco de programas de *compliance* de empresas privadas norteamericanas<sup>65</sup> y que ya hemos avanzado en el subapartado anterior: el recordatorio moral o ético, en forma de mensaje, antes de realizar alguna actividad sujeta a riesgos de incumplimiento legal o ético, como podría ser en el momento de llenar los formularios para el reembolso de gastos de viaje (en cuyo caso, tiene un impacto comprobado en la reducción de irregularidades cuando dicho recordatorio se encuentra en el encabezamiento del formulario).

Este tipo de medida de *nudging*, en nuestra opinión, sería relativamente sencilla de poner en práctica y podría aplicarse, adaptándose a cada ámbito, en distintas actividades administrativas en las que se haya detectado un número elevado de vulnerabilidades o un nivel de riesgo significativo. Así, por ejemplo, en una actuación administrativa donde se haya detectado riesgo de conflicto de intereses, se podría poner a disposición de la persona titular del órgano competente para adoptar la resolución correspondiente un mensaje recordando la prohibición de actuar en caso de concurrir algún conflicto de intereses (listando las causas de abstención, en su caso) y recordando la

64 OCDE, *Behavioural Insights for Public Integrity. Harnessing the Human Factor to Counter Corruption*, 2018.

65 TODD HAUGH, "Nudging Corporate Compliance", *op. cit.*, pp. 710 y ss.

obligación, en caso de que así sea, de declarar cualquier información pertinente para mantener actualizados los registros públicos de actividades y bienes e intereses. Siguiendo con técnicas de *nudging*, incluso se podría añadir el número de personas que, cumpliendo con sus obligaciones de declaración y actualización de los datos en tales registros<sup>66</sup>, han comprobado, y en su caso actualizado, sus datos en los últimos tres meses, por ejemplo.

Podemos, también, pensar fácilmente en otro posible ejemplo a aplicar en el sector público en relación con el cumplimiento de los plazos para resolver los procedimientos administrativos. Así, sería posible que cuando se detecte un retraso en la tramitación de un expediente, unos días antes de la finalización del plazo para resolver, se advierta al órgano responsable de la tramitación del procedimiento del tiempo restante para resolver el procedimiento, indicando también del tiempo promedio en que se resuelven ese tipo de expedientes en aquella administración (siempre que, dicho tiempo promedio, sea igual o inferior al plazo máximo legalmente establecido para resolver ese procedimiento). También se podría pensar en añadir cargas adicionales en caso de incumplimiento de los plazos, como por ejemplo podría ser la petición, a los órganos responsables de la tramitación de los procedimientos, de motivar las demoras en el momento de vencer los plazos.

Por último, debemos señalar que este tipo de *nudges* en el sector público, que buscarán asegurar el cumplimiento normativo y la conducta ética, tenderán a ser diseñados para entornos digitales en la medida en que progresivamente se va implantando la administración electrónica<sup>67</sup>. Serán, por tanto, *nudges* digitales en el funcionamiento electrónico de la Administración pública, y también *hyper nudges*, cuando estos se apoyen en tecnologías basadas en inteligencia artificial y en los análisis de *big data* con el fin de ofrecer acicates o *nudges* éticos especialmente adaptados a las circunstancias del caso y particularmente dirigidos a aquellos sujetos que, por la concurrencia de significativos factores de riesgo, presentan mayor probabilidad de tomar decisiones contrarias a la ética en una actuación determinada<sup>68</sup>.

66 Por ejemplo, Ley 7/1985, del 2 de abril, artículo 75.7, Reguladora de las Bases del Régimen Local.

67 Así, el artículo 70.2 de la Ley 39/2015 impone el formato electrónico para los expedientes administrativos ("Los expedientes tendrán formato electrónico y se formarán mediante la agregación ordenada de cuantos documentos, pruebas, dictámenes, informes, acuerdos, notificaciones y demás diligencias deban integrarlos, así como un índice numerado de todos los documentos que contenga cuando se remita. Asimismo, deberá constar en el expediente copia electrónica certificada de la resolución adoptada"), y el artículo 36 de la misma ley fija también, con carácter general, que los actos administrativos "se producirán por escrito a través de medios electrónicos, a menos que su naturaleza exija otra forma más adecuada de expresión y constancia".

68 En este sentido, Feldman y Kaplan sostienen que: "By utilizing big data analytics, policy makers can predict unethicability and then deploy targeted regulatory responses in real time in order to

La práctica administrativa en un campo conexo con la prevención de las malas prácticas y la corrupción, como es la lucha contra el fraude fiscal, ofrece un ejemplo de *hyper nudging* que fácilmente podría trasladarse al funcionamiento interno de las administraciones para luchar contra las malas prácticas de la propia Administración pública. Se trata de un sistema desarrollado recientemente por la Agencia Tributaria española para la campaña de la declaración anual de la renta basado en el *nudging* y el análisis de *big data* mediante inteligencia artificial<sup>69</sup>. El sistema tiene capacidad para predecir errores en la confección de las autoliquidaciones del impuesto sobre la renta de las personas físicas, y está preparado para presentar al contribuyente un mensaje de alerta advirtiéndole de que hay una alta probabilidad de que su declaración sea incorrecta, cuando algún dato aportado o modificado por el contribuyente no coincide con los datos disponibles en la Agencia Tributaria. Asimismo, esta herramienta también estaría preparada para el control de las autoliquidaciones, identificando aquellos contribuyentes con elevado riesgo de haber presentado una declaración errónea modificando los datos a disposición de la Agencia Tributaria.

#### 4. ALGUNOS RIESGOS DERIVADOS DEL EMPLEO DE NUDGES EN ENTORNOS DIGITALES

Veamos, a continuación, algunos de los riesgos generalmente asociados a los *nudges*, especialmente en entornos digitales, para después concretar cómo estos se pueden llegar a dar en el sector público, en el diseño y empleo de herramientas antifraude y anticorrupción.

##### 4.1. LOS DISEÑOS OSCUROS (DARK PATTERNS)

Los *diseños oscuros* podrían definirse como diseños de interfaz de usuario y experiencia de usuarios que intentan intencionadamente explotar las vulnerabilidades de las personas, manipulando y engañando a los usuarios<sup>70</sup>. Este

*improve ethical deliberation by potential perpetrators. The use of targeted ethical reminders can improve ethical decision-making, without running the risk of creating ethical numbing and over bombarding people with meaningless and random ethical messages. Big data analysis can also help tailor the most appropriate regulatory response to each specific case and match it to the specific bias that is creating the opportunity for unethical conduct". FELDMAN y YOTAM KAPLAN, "Big Data and bounded ethicality", op. cit., p. 93.*

- 69 Ver Resolución del 19 de enero de 2021, de la Dirección General de la Agencia Estatal de Administración Tributaria, por la que se aprueban las directrices generales del Plan Anual de Control Tributario y Aduanero de 2021.
- 70 Por todos, NORWEGIAN CONSUMER COUNCIL, *Deceived by Design: How Tech Companies Use Dark Patterns to Discourage Us From Exercising Our Rights to Privacy*, Forbrukerradet, 2018.

tipo de diseños puede utilizarse con diversos fines (por ejemplo, obtener más datos personales, dinero o influir en un voto)<sup>71</sup>.

Incluso el derecho positivo se ha referido a este tipo de diseños. Así, la California Privacy Rights Act de 2020 se refiere a estos diseños oscuros en los siguientes términos: "*'Dark pattern' means a user interface designed or manipulated with the substantial effect of subverting or impairing user autonomy, decision making, or choice, as further defined by regulation*"<sup>72</sup>.

Los diseños oscuros<sup>73</sup> buscan utilizar los sesgos de los consumidores de una forma deshonesta para empujarlos a cometer errores o lograr o evitar determinadas actuaciones, mediando técnicas de manipulación basadas en conocimientos derivados de las ciencias del comportamiento. En efecto, los diseños oscuros están hechos para aprovechar los sesgos cognitivos, esto es, los errores sistemáticos creados por atajos mentales (heurísticos, que como humanos y consumidores nos afectan en nuestra toma de decisiones, y también para usar las emociones humanas en favor de las empresas privadas que actúan como arquitectas del marco de elección).

En la literatura científica es posible identificar ya estudios que relacionan dichos diseños oscuros con los sesgos cognitivos que manipulan en favor de las empresas<sup>74</sup>.

Procede, ante esta situación, plantearse cómo deben los poderes públicos abordar este empleo negativo de las ciencias conductuales en los diseños de entornos decisionales. Desde luego, puede existir diversos tipos de intervención, como podría ser mediante la actividad de policía y fomento, mediante campañas de información o fomentando la autorregulación de las empresas (como está impulsando la Unión Europea, mediante su Código de Prácticas de Desinformación de 2018), o incluso alcanzando hasta la plena regulación jurídica de la arquitectura de la decisión digital, esto es, prohibiendo y estableciendo determinados requisitos de protección del consumidor, con tipificación en su caso de infracciones y sanciones<sup>75</sup>.

71 Esta idea se encuentra basada en el proyecto de investigación PID2020-115774RB-I00 ya mencionado.

72 Puede traducirse por: "Diseño oscuro: una interfaz de usuario diseñada o manipulada con el efecto sustancial de subvertir o perjudicar la autonomía la toma de decisiones o la elección del usuario, tal y como se define en la normativa".

73 Sobre diseños oscuros y algunos ejemplos de estos, véase, por todos, JAMIE LUGURI y LIOR JACOB STRAHILEVITZ, "Shining a Light on Dark Patterns", *Journal of Legal Analysis*, vol. 13, n.º 1, 2021, pp. 43-109.

74 ARUNESH MATHUR *et al.*, "Dark patterns at scale: Findings from a crawl of 11K shopping websites", en *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 2019, pp. 1-32.; y SUSAN WEINSCHENK, *Diseño inteligente. Las 100 cosas que todo diseñador necesita saber sobre las personas*, Madrid: Anaya Multimedia, 2021.

75 Esta idea se encuentra basada en el proyecto de investigación PID2020-115774RB-I00 ya mencionado.

De hecho, tanto la Unión Europea como los Estados Unidos cuentan con regulaciones de los diseños oscuros, como se ha apuntado, pensadas principalmente para las relaciones entre usuarios o clientes y empresas. Cabría añadir, no obstante, si esa regulación sería también suficiente para cubrir esos supuestos en los que es la propia Administración la que emplea un sistema de control de las actividades de su personal propio (los servidores públicos), en el que podría haber algún diseño oscuro que busque condicionar de forma excesiva la voluntad de los servidores públicos en el desarrollo de sus funciones, explotando los sesgos antes referidos.

#### 4.2. RIESGOS ESPECÍFICOS DE LOS NUDGES DE PRECISIÓN (*HYPERNUDGING*)

En el contexto actual no podemos, sin embargo, separar el empleo de *nudges* del uso de herramientas digitales. En esta línea se desarrolla lo que Yeung ha denominado *hypernudge*<sup>76</sup>, entendido como un *nudge* potenciado mediante el uso del *big data* y los algoritmos. En efecto, estas herramientas digitales presentarían una capacidad extraordinaria para el diseño de *nudges* "a medida", de precisión, mediante el aprendizaje automático de los sistemas más avanzados basados en inteligencia artificial.

Yeung ha destacado que el *nudging* impulsado por *big data* es capaz de generar una arquitectura de elección altamente personalizada, *guiando* las decisiones del usuario, sea consumidor ante una empresa o sea el usuario de un servicio público, de ahí que esta autora se refiera a estas técnicas como "*hypernudge*".

Las herramientas digitales dedicadas al análisis de datos para personalizar servicios pueden, en efecto, resultar beneficiosas. Por ejemplo, en el ámbito sanitario, la pandemia del COVID-19 ha dado lugar al uso en diversos países de la tecnología como un instrumento para hacerle frente. La tecnología ha sido empleada para diversos fines, como la consecución de una vacuna, el intercambio de conocimientos, la observación y predicción de la pandemia y la ayuda al personal sanitario para el diagnóstico del coronavirus, o como herramienta para el control de la extensión de la pandemia<sup>77</sup>.

Centrándonos ahora en la última de las finalidades, diversos países, como Singapur, Israel, Taiwán, Corea del Sur, Islandia o Australia, por ejemplo, han diseñado sistemas tecnológicos del control de los riesgos epidémicos. Estos sistemas pueden ir desde la geolocalización de los móviles por los operadores de telecomunicaciones o a partir de las redes sociales hasta las cámaras de

76 KAREN YEUNG, "'Hypernudge': Big Data as a mode of regulation by design", *Information, Communication & Society*, vol. 20, n.º 1, 2017, pp. 118-136.

77 CONSEJO DE EUROPA, *Towards regulation of ai systems*, op. cit.

infrarrojos para lecturas masivas de temperatura, pasando por aplicaciones (apps) para autotest o cita previa (con también webs o *chatbots* para la misma finalidad), apps de información voluntaria de contagios, apps de seguimiento de contactos por Bluetooth (*contact trace apps*) y pasaportes de inmunidad<sup>78</sup>.

En cuanto a las dificultades de aplicación, un ejemplo paradigmático lo presentan las apps voluntarias de rastreo de contagio. En general, no han sido excesivamente útiles hasta el momento de escribir estas líneas, debido a un problema fundamental para el éxito de estas tecnologías: para que funcionen eficazmente deben ser adoptadas por un número suficiente de personas (hablamos de muchos millones de personas).

En las experiencias existentes hasta el momento, como Australia o Islandia, el uso de las apps es alto en los primeros momentos, pues los ciudadanos, por curiosidad o por motivación personal de protección a los conciudadanos, se las descargan masivamente al principio. Sin embargo, conforme pasa el tiempo, las descargas decrecen y, por otro lado, dejan de ser usadas.

Una solución pudiera ser establecer legislativamente la obligación del uso de dichas apps con las correspondientes sanciones por incumplimiento (lo que debería hacerse por norma con rango de ley de forma clara y adecuada<sup>79</sup>). Ahora bien, lo que nos interesa en este trabajo es una alternativa para incrementar el uso de las apps no basada en el ordeno y mando, en la obligación y el castigo, sino en incentivos basados en los conocimientos que nos proporcionan las ciencias del comportamiento y que podrían ser utilizados por las AAPP.

Nos referimos al diseño de acicates para la descarga y uso de dichas apps. Conociendo los sesgos cognitivos de la población, podrían diseñarse soluciones para romper el sesgo de inercia. Como que la descarga de la app ofrezca información atractivamente diseñada sobre la pandemia a quien lo haga o que quien se descargue y use la app pueda entrar en una lotería del sector público para recibir un premio en metálico, que podría financiarse usando las sanciones impuestas a quienes infrinjan las normas de prevención de contagio de la pandemia. La ciudad de Estocolmo ha diseñado una lotería semejante para aquellos que respeten las normas de tráfico, financiada con

78 AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS, *El uso de las tecnologías en la lucha contra el COVID-19. Un análisis de costes y beneficios*, Madrid: Agencia Española de Protección de Datos, 2020; y CONSEJO DE EUROPA, *Towards regulation of ai systems*, op. cit.

79 LORENZO COTINO, "Inteligencia artificial, big data y aplicaciones contra la COVID-19: privacidad y protección de datos", *Revista de Internet, Derecho y Política*, n.º 31, octubre, 2020, pp. 1 y ss. Señala este autor como "La eficacia de estas herramientas dependerá de su uso masivo y, si no lo hay voluntariamente, no hay que descartar su obligatoriedad sobre la base de una clara legalidad" (p. 11).

las multas impuestas a los malos conductores<sup>80</sup>. Incluso se ha planteado una opción por defecto en favor de la descarga<sup>81</sup>.

En cuanto a los riesgos de este acicate en el uso de estas apps, queremos destacar aquí dos: el tratamiento ilegítimo de datos personales<sup>82</sup> y la manipulación de los ciudadanos, siendo este segundo el que ahora nos interesa. El posible salto de la personalización a la manipulación debiera ser evitado, pues alteraría el principio jurídico de buena fe en que deben basarse las relaciones entre Administraciones y ciudadanos, y supondría también en términos jurídicos una desviación de poder: esto es, el uso de una potestad otorgada con una finalidad de interés general concreta (recogida y tratamiento de datos personales sanitarios para luchar contra una pandemia, en su caso mediante

- 80 En el caso de Estocolmo, el municipio sueco diseño un sistema por el que cuando un coche pasaba por un cruce específico, una cámara tomaba una foto de este y medía su velocidad. Si el conductor superaba el límite de velocidad, se le imponía una multa. El dinero se usaba para enriquecer un fondo único, el premio de la lotería para conductores obedientes. *The Medical Futurist*, "The Swedish Speed Camera Lottery and Healthy Living", 2018. Disponible en: <https://medicalfuturist.com/swedish-speed-camera-lottery-healthy-living/>. Más en general, sobre el problema de las mencionadas apps, véase LUIS MILLER, "¿Qué esperar de las aplicaciones de rastreo de contagios?", *elDiario.es*, 9 de junio de 2020. Disponible en: [https://www.eldiario.es/piedrasdepapel/aplicaciones-rastreo-contagios-covid19\\_132\\_6023027.html](https://www.eldiario.es/piedrasdepapel/aplicaciones-rastreo-contagios-covid19_132_6023027.html)
- 81 MICHELLE MELLO y JASON WANG, "Ethics and governance for digital disease surveillance". *Science*, vol. 368, n.º 6494, 2020, p. 951-954.
- 82 En cuanto al *tratamiento de datos personales*, hay que empezar dejando claro que la normativa de protección de datos, al menos la europea y la española, permite el tratamiento legítimo de datos personales en situaciones como la pandemia de COVID-19 en la que existe una emergencia sanitaria. Permiten dicho tratamiento de datos personales el Reglamento UE 2016/679 del Parlamento Europeo y del Consejo del 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (considerando 46, artículos 6.1c, cumplimiento de una misión legal, 6.1 d, intereses vitales del interesado u otras personas físicas, 6.1 e, misión realizada en interés público, 9.2, por ejemplo, letras c, g, h, o i, la ley orgánica española 3/1986, del 14 de abril, de Medidas Especiales en Materia de Salud Pública (modificada por el Real Decreto-Ley 6/2020, del 10 de marzo, artículo 3), la Ley española 33/2011, del 4 de octubre, General de Salud Pública (artículos 5 y 84). Ahora bien, el mencionado tratamiento deberá realizarse siempre respetando el Reglamento UE aludido (incluyendo sus principios del artículo 5.º) y la normativa española (Ley orgánica 3/2018, del 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales: téngase en cuenta la Disposición Adicional decimoséptima), de acuerdo con el aludido principio de proporcionalidad y sin que los datos sean tratados por terceros con otros fines (considerando 54 del Reglamento UE). Sobre detalles técnicos, véase LORENZO COTINO, "Inteligencia artificial, big data y aplicaciones contra la COVID-19", *op. cit.*, pp. 12 y 13. En el caso español, en julio de 2020 se lanzó en prueba Radar COVID, la aplicación diseñada y dirigida por la Secretaría de Estado de Digitalización e Inteligencia Artificial del Gobierno de España para ayudar a evitar la propagación del coronavirus (COVID-19).

apps) para otros fines distintos (manipulación del comportamiento de los ciudadanos, por ejemplo, persuadiéndoles para no utilizar el servicio público sanitario o para no ejercer derechos constitucionalmente garantizados).

En segundo lugar, en ámbitos distintos al sanitario, destaca el caso del *hypernudge* desarrollado por la Agencia Tributaria española en ayuda del contribuyente<sup>83</sup>, al que ya hemos hecho una breve mención anteriormente. El sistema tiene capacidad para predecir errores en la confección de las autoliquidaciones del impuesto sobre la renta de las personas físicas, y está preparado para presentar al contribuyente un mensaje de alerta advirtiéndole de que hay una alta probabilidad de que su declaración sea incorrecta, cuando algún dato aportado o modificado por el contribuyente no coincide con los datos disponibles en la Agencia Tributaria. Asimismo, esta herramienta también estaría preparada para el control de las autoliquidaciones, identificando aquellos contribuyentes con elevado riesgo de haber presentado una declaración errónea modificando los datos a disposición de la Agencia Tributaria<sup>84</sup>. De nuevo, una ayuda personalizada a los contribuyentes no debería degenerar en ningún caso en una manipulación.

Puede destacarse, también, otro ejemplo en el que, una herramienta con capacidad para incidir en el control del cumplimiento normativo, ha sido considerada generadora de excesivos riesgos desde la perspectiva de la protección de datos personales. Se trata del caso de la herramienta SyRI, anulada por sentencia del Tribunal de Distrito de La Haya (Países Bajos), del 5 de febrero de 2020. Empleando esta herramienta, las autoridades holandesas recogieron durante años datos sobre ingresos, pensiones, seguros, tipo de casa, impuestos, multas, integración, educación, deudas o subsidio de desempleo de los contribuyentes para elaborar perfiles y calcular luego mediante algoritmos las probabilidades de los sujetos tratados por el programa de defraudar a la Administración. La decisión judicial encontró desproporcionada dicha recogida de datos, en relación con la opacidad del sistema y la ausencia de suficientes garantías de control de su correcto funcionamiento y supervisión<sup>85</sup>.

83 Este *nudge* de precisión es explicado muy bien en este video, FUNDACIÓN COMPLUTENSE, "Fiscalidad e Inteligencia Artificial: Sistema fiscal y procedimientos tributarios (primera sesión)", Congreso Internacional Virtual, 2021. Disponible en: <https://www.youtube.com/watch?v=9ynftKLBBK8>, desde la franja de tiempo de 13 minutos de este, que va desde las 2 horas y 8 minutos hasta las 2 horas y 21 minutos.

84 Ver Resolución del 19 de enero de 2021, de la Dirección General de la Agencia Estatal de Administración Tributaria, por la que se aprueban las directrices generales del Plan Anual de Control Tributario y Aduanero de 2021.

85 Esta idea se encuentra basada en el proyecto de investigación PID2020-115774RB-I00 ya mencionado.

#### 4.3. LA CONCRECIÓN DE LOS RIESGOS GENERALES EN LA FUNCIÓN DE CONTROL EN EL SERVICIO PÚBLICO

El primer riesgo que destacar es que este tipo de programas podrían llegar a almacenar una gran cantidad de datos sobre los servidores públicos, a fin de reducir y controlar sus riesgos de corrupción, y sin embargo, no parece que en la actualidad se estén estableciendo mecanismos suficientes para compensar la intromisión de estos programas o sistemas en la vida de las personas. Así, si una persona (un servidor público) que fuera objeto de tratamiento de sus datos por parte del programa anticorrupción quisiera conocer cómo funciona el programa y exactamente para qué emplea sus datos, es posible con que encuentre restricciones. Por ejemplo, si nos centramos ahora en la normativa española, además del límite general que se podría atribuir a todos los sistemas informáticos, consistente en la protección de la propiedad intelectual e industrial (artículo 14.1.j de la Ley estatal 19/2013), en los casos de los sistemas antifraude y anticorrupción, que son, en general, sistemas para prevenir y detectar prácticas ilícitas, existiría un límite adicional, previsto en el artículo 14.1.e de la misma Ley 19/2013, según el cual se puede limitar el derecho de acceso cuando suponga un perjuicio para la "prevención, investigación y sanción de los ilícitos penales, administrativos o disciplinarios".

En efecto, no es extraño que, con el legítimo fin de luchar contra el fraude y la corrupción, algunos de estos programas pueden acceder a datos personales (véase, por ejemplo, la posibilidad de acceso a datos personales del sistema de alertas valenciano, artículo 17.3 de la Ley valenciana 22/2018). En este punto, otra pregunta que se puede formular es si la intromisión en la privacidad guarda una relación de proporcionalidad adecuada, o si se trata de un sistema sin suficientes garantías de control de su correcto funcionamiento y supervisión, como se puede observar en el ya referido caso resuelto en la sentencia del Tribunal de Distrito de La Haya (Países Bajos) sobre el sistema SyRI, del 5 de febrero de 2020.

En segundo lugar, debemos considerar el riesgo de que el sistema inteligente anticorrupción, en aplicación de medidas preventivas o de mitigación de riesgos de corrupción, plantee entornos decisionales que, mediante manipulación, conduzcan de forma especialmente intensa hacia una determinada solución, siendo esta la que el programa considera la que presenta un más bajo riesgo de corrupción. Como punto de partida, debemos recordar que la toma de decisiones pública debe basarse en una informada y adecuada ponderación de los distintos intereses en juego, tanto públicos como privados, para así adoptar la opción que se considere más adecuada a los intereses generales. Por tanto, la manipulación del decisor público, orientándolo hacia una solución o actuación específica, puede conducir en último término a un análisis sesgado, incompleto de los asuntos, y a desvirtuar los ámbitos en los que se ha considerado necesaria la ponderación de intereses de acuerdo

con las circunstancias de cada momento, como son los ámbitos donde se reconocen potestades discrecionales<sup>86</sup>.

Debemos recordar, en este punto, que el empleado público accede a su puesto demostrando su mérito y capacidad en unas pruebas de acceso específicas, que acreditan *a priori* su idoneidad para desempeñar correctamente su función, mientras que, por el momento, no parece existir un sistema equivalente para verificar la idoneidad del programa que se quiera emplear para actuar, de forma explícita o implícita, en sustitución del criterio del empleado público. Y tampoco parece que el diseño de este tipo de programas o sistemas responda a una representación democrática, como se daría en los cargos electos que también adoptan decisiones en el entramado institucional de gobiernos y administraciones públicas.

## CONCLUSIONES: ALGUNAS REFLEXIONES FINALES SOBRE NUDGING E INTELIGENCIA ARTIFICIAL

En definitiva, la combinación de diseño digital, *nudges*, explotación de sesgos cognitivos, *big data* e IA puede ayudar a mejorar el funcionamiento del mercado y la gestión pública, incluyendo un espacio significativo para mejorar la actuación ética y anticorrupción en el sector público, pero también puede crear efectos adversos.

En el ámbito europeo, la futura regulación por la Unión Europea de la inteligencia artificial (IA) aparece como una ocasión idónea para debatir también estos temas, yendo más allá de la preocupación, necesaria pero insuficiente, por la protección de datos personales. En la propuesta de regulación europea<sup>87</sup> se establece una lista de IA prohibidas. La regulación propuesta sigue un enfoque basado en el riesgo diferenciando entre los usos de IA en función de los riesgos que llevan, clasificándolos en sistemas que crean: (1) un riesgo inaceptable, (2) un riesgo alto y (3) riesgo bajo o mínimo. La lista de prácticas prohibidas del Título II comprende todos aquellos sistemas de IA cuyo uso se considera inaceptable por contravenir los valores de la Unión, por ejemplo, "la comercialización, puesta en servicio o utilización de un sistema de IA que despliegue técnicas subliminales más allá de la conciencia

86 Sobre inteligencia artificial y discrecionalidad, resultan de interés las reflexiones de ALEJANDRO HUERGO, "Administraciones públicas e inteligencia artificial: ¿Más o menos discrecionalidad?", *El Cronista del Estado Social y Democrático de Derecho*, n.º 96-97, 2021, pp. 78-95.

87 Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley De Inteligencia Artificial) y se modifican determinados actos legislativos de la unión, COM/2021/206 final, consultable en: <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:52021PC0206>.

de una persona para distorsionar materialmente el comportamiento de una persona de manera que cause o pueda causar a esa persona o a otra un daño físico o psicológico<sup>88</sup>.

Téngase en cuenta que la propuesta señala que otras prácticas de manipulación o explotación que afectan a adultos que podrían ser facilitadas por los sistemas de IA podrían estar cubiertas por la protección de datos, protección de los consumidores y legislación sobre servicios digitales que garantizan que las personas físicas sean debidamente informadas y tengan la libre elección de no ser sometidas a la elaboración de perfiles u otras prácticas que puedan afectar a su comportamiento.

Sin embargo, dada la relevancia de los diseños digitales con IA potencialmente manipulativos, creemos que *la alusión a daño físico o psicológico no es demasiado afortunada* y sugerimos que debiera substituirse por una simple referencia a la posibilidad de causar o inducir a error o engaño, afectando a la autonomía de la voluntad. En este sentido, ya aludimos a la reciente modificación de la legislación californiana, la California Privacy Rights Act (CPRA) de 2020, para prohibir los diseños oscuros.

En el ámbito español, por ejemplo, se está impulsando desde hace tiempo una Carta de Derechos Digitales<sup>89</sup>. Estas cuestiones aquí tratadas creemos que debieran estar presentes en la versión que finalmente se apruebe. En el texto sometido a consulta pública se indica que "Se prohíbe el uso de sistemas de inteligencia artificial dirigidos a manipular o perturbar la voluntad de las personas, en cualesquiera aspectos que afecten a los derechos fundamentales" y más específicamente respecto a los menores se señala que "Se consideran ilícitas las prácticas de perfilado susceptibles de manipular o perturbar la voluntad de los menores y, en particular, la publicidad basada en este tipo de técnicas". En esta línea, sería interesante avanzar en la precisión de la "manipulación o perturbación", acudiendo a los ejemplos norteamericano y de la UE vistos, en los que se acude a los conceptos de causación de error o engaño por el diseño digital, incluyendo el uso de técnicas subliminales, que se dirigen al sistema de pensamiento automático, para inducir a ese error o engaño.

Ahora bien, junto a estos preocupantes riesgos de engaño y manipulación, hemos destacado también como el uso de *nudges* digitales en combinación con la IA puede suponer para el sector público un impulso de la buena

88 Esta idea se encuentra basada en el proyecto de investigación PID2020-115774RB-I00 ya mencionado.

89 MINISTERIO DE ASUNTOS ECONÓMICOS Y TRANSFORMACIÓN DIGITAL, *Consulta pública para la elaboración de una carta de derechos digitales*, Madrid: Secretaría de Estado de Digitalización e Inteligencia Artificial, 2020. Disponible en: [https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/participacion\\_publica/consulta/ficheros/Carta-DerechosDigitales.pdf](https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/participacion_publica/consulta/ficheros/Carta-DerechosDigitales.pdf).

administración, mediante la mejora de la calidad normativa, de la transparencia, de la personalización de los servicios públicos, como se ha comprobado durante la pandemia, y aspecto que hemos analizado de forma central en este estudio, para la prevención de la corrupción y la mala administración. Todos estos aspectos van a cobrar, si cabe, aún más relevancia en la situación pospandemia, y deberá prestarse especial atención al diseño de mecanismos de garantía para evitar que este tipo de sistemas digitales incurran en algunos de los riesgos y problemas que se han ido exponiendo a lo largo del trabajo, como la manipulación, la recopilación de datos excesivos o la opacidad.

El reto que se presenta, pues, no es menor: habrá que encontrar en la gestión privada y en la gestión pública, y en el derecho que regula a los sectores privado y público, el equilibrio preciso para el servicio adecuado a los intereses generales.

## BIBLIOGRAFÍA

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. *El uso de las tecnologías en la lucha contra el COVID-19. Un análisis de costes y beneficios*. Madrid: Agencia Española de Protección de Datos, 2020.

AMOEDO, JOSÉ DANIEL. "Los sistemas informáticos de detección de malas prácticas: herramientas esenciales para la prevención de la corrupción". *Revista Internacional Transparencia e Integridad*, n.º 6, 2018.

BARENDs, ERIC et al. *Unethical behaviour in the workplace: A rapid assessment of the research literature*. CIPD: Technical Report, 2019.

BARRAGUÉ, BORJA, Y JOSÉ NOGUERA. "Los nudges y el diseño conductual de políticas públicas". *Gestión y Análisis de Política Pública*, n.º 25, 2021.

BEN-ISRAEL, ISAAC, JORGE CERDIO et al. *Towards Regulation of AI Systems Global perspectives on the development of a legal framework on Artificial Intelligence (AI) systems based on the Council of Europe's standards on human rights, democracy and the rule of law*, 2020.

CABRALES, ANTONIO, Y PEDRO REY BIEL. "Mas allá de los nudges: Políticas públicas efectivas basadas en la evidencia de las ciencias del comportamiento". GAPP, n.º 25, 2021, pp. 38-45.

CAMPOS ACUÑA, CONCEPCIÓN. *Aplicación práctica del compliance en la contratación pública*. Pamplona: Aranzadi, 2019.

CAMPOS ACUÑA, CONCEPCIÓN. *Compliance en la Administración pública: Dificultades y propuestas*. Granada: CEMCI, 2017.

CAMPOS ACUÑA, CONCEPCIÓN. *Guía práctica de compliance en el sector público*. Madrid: Walters Kluwer, 2020.

CAPDEFERRO, ÓSCAR. "Las herramientas inteligentes anticorrupción: entre la aventura tecnológica y el orden jurídico". *Revista General de Derecho Administrativo*, n.º 50, 2019.

CASSESE, SABINO. "Exploring the legitimacy of Nudging". En Alexandra Kemmerer, Christoph Möllers, Maximilian Steinbeis y Gerhard Wagner (eds.), *Choice Architecture in Democracies. Exploring the Legitimacy of Nudging*. Baden-Baden: Nomos, 2016.

CASSESE, SABINO. "New Paths for Administrative Law: A Manifesto". *International Journal of Constitutional Law*, vol. 10, n.º 3, 2012.

CERRILLO, AGUSTÍ, Y MIQUEL PEGUERA (coords.). *Retos jurídicos de la inteligencia artificial*. Cizur Menor: Aranzadi, 2020.

COMISIÓN EUROPEA. *A Handbook for Citizen-centric eGovernment*. Ecotec, 2007.

CONSEJO DE EUROPA. "Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes", *Committee of Ministers*, 13 February 2019, 1337 meeting of the Ministers' Deputies, 2019.

CONSEJO DE EUROPA. *Towards Regulation of AI Systems Global perspectives on the development of a legal framework on Artificial Intelligence (AI) systems based on the Council of Europe's standards on human rights, democracy and the rule of law*, 2020.

COSTAS, ELENA, Y PABLO TUCAT. "Nudges: diseño y evaluación", *GAPP*, n.º 25, 2021, pp. 8-22.

COTINO, LORENZO. "Ética en el diseño para el desarrollo de una inteligencia artificial, robótica". *Revista catalana de dret públic*, n.º 58, 2019.

COTINO, LORENZO. "Inteligencia artificial, big data y aplicaciones contra la COVID-19: privacidad y protección de datos". *Revista de Internet, Derecho y Política*, n.º 31, 2020.

DE GRAAF, GJALT, LEO HUBERTS Y TEBBINE STRÜWER. "Integrity violations and corruption in Western public governance: Empirical evidence and reflection from the Netherlands", *Public Integrity*, vol. 20, n.º 2, 2018.

FELDMAN, YUVAL, Y YOTAM KAPLAN. "Big Data and bounded ethicality". *Cornell Journal of Law and Public Policy*, vol. 29, n.º 39-93, 2019.

FRÍAS, AMPARO. "Los servicios personalizados de información de actualidad de los medios de comunicación españoles a través de Internet". *BiD*, n.º 17, 2006.

- GRANDE SERRANO, PABLO. "La psicología económica como herramienta para incentivar el cumplimiento voluntario de las obligaciones tributarias". *Revista de Contabilidad y Tributación*, n.º 439, 2019, pp. 5-36.
- HAUGH, TODD. "Nudging Corporate Compliance". *American Business Law Journal*, vol. 54, n.º 4, 2017.
- HAUSER, CHRISTIAN. "Fighting Against Corruption: Does Anti-corruption Training Make Any Difference?". *Journal of Business Ethics*, vol. 159, n.º 1, 2019.
- HAUSER, CHRISTIAN. "From Preaching to Behavioral Change: Fostering Ethics and Compliance Learning in the Workplace". *Journal of Business Ethics*, vol. 162, n.º 4, 2020.
- HUERGO, ALEJANDRO. "Administraciones públicas e inteligencia artificial: ¿O menos discrecionalidad?". *El Cronista del Estado Social y Democrático de Derecho*, n.º 96-97, 2021.
- KAHNEMAN, DANIEL. *Pensar rápido, pensar despacio*. Barcelona: Debolsillo, 2013.
- LANGEVOORT, DONALD. "Behavioral Ethics, Behavioral Compliance". En Jennifer Arlen, *Research Handbook on Corporate Crime and Financial Misdealing*. Nueva York: Edward Elgar Publishing, 2018.
- LUGURI, JAMIE, Y LIOR JACOB STRAHILEVITZ. "Shining a Light on Dark Patterns". *Journal of Legal Analysis*, vol. 13, n.º 1, 2021, pp. 43-109.
- MATHUR, ARUNESH, *et al.*, "Dark patterns at scale: Findings from a crawl of 11K shopping websites". *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 2019.
- MELLO, MICHELLE, Y JASON WANG. "Ethics and governance for digital disease surveillance". *Science*, vol. 368, n.º 6494, 2020.
- MIRANZO DÍAZ, JAVIER. "Inteligencia artificial y contratación pública". En Ignacio Martín y José Antonio Moreno Molina (dirs.), *Administración electrónica, transparencia y contratación pública*. Madrid: Iustel, 2020.
- NIETO MARTÍN, ADÁN. "De la ética pública al *public compliance*: Sobre la prevención de la corrupción en las administraciones públicas". En Adán Nieto Martín (dir.), *Public compliance. Prevención de la corrupción en Administraciones públicas y partidos políticos*. Madrid: UCLM, 2014.
- NORWEGIAN CONSUMER COUNCIL. *Deceived by Design: How Tech Companies Use Dark Patterns to Discourage us from Exercising Our Rights to Privacy*. Forbrukerradet, 2018.

OCDE. *Behavioural Insights for Public Integrity. Harnessing the Human Factor to Counter Corruption*, 2018.

OCDE. *Recomendación del Consejo de la OCDE sobre integridad pública*. OCDE, 2017.

OCDE. *Tools and Ethics for Applied Behavioural Insights: The BASIC Toolkit*. OECD Publishing, 2019.

OCDE. *Towards a Sound Integrity Framework: Instruments, Processes, Structures and Conditions for Implementation*, GOV/PGC/GF(2009)1, OCDE, 2009

PONCE, JULI. "La lucha por el buen gobierno y el derecho a una buena administración mediante el estándar jurídico de diligencia debida". *Cuadernos Democracia y Derechos Humanos*, n.º 15, 2019.

PONCE, JULI. "La prevención de riesgos de mala administración y corrupción, la inteligencia artificial y el derecho a una buena administración". *Revista Internacional de Transparencia e Integridad*, n.º 6, 2018.

PUNCEL CHORNET, ALFONSO. "Inteligencia artificial para la transparencia pública: El Sistema de Alertas Tempranas (Saler) de la Generalitat Valenciana". *Boletín Económico de ICE*, n.º 3116, 2019.

QUERALT, JOAN. "Public compliance y corrupción: Análisis conceptual y propuestas". *Revista Internacional de Transparencia e Integridad*, n.º 2, 2016.

RANCHORDÁS, SOFÍA. "Nudging citizens through technologies in Smart Cities". *International Review of Law, Technology and Computers*, n.º 33, 2019.

RIEGER, SEBASTIAN, Y CAROLINE SINDERS. *Dark Patterns: Regulating Digital Design*. Stiftung Neue Werantwortung, 2020.

SUNSTEIN, CASS, Y RICHARD TAHLER. *Un pequeño empujón (nudge)*. Madrid: Taurus, 2009.

TEJADA PLANA, DANIEL. *Behavioral Compliance: Reforzando el compliance a través de la ética conductual*. Cizur Menor: Thomson Reuters Aranzadi, 2021.

THALER, JULIA, Y BERND HELMIG. "Do codes of conduct and ethical leadership influence public employees' attitudes and behaviours? An experimental analysis". *Public Management Review*, vol. 18, n.º 9, 2016.

THALER, RICHARD H., CASS R. SUNSTEIN, Y JOHN P. BALZ. "Choice Architecture". En *The Behavioral Foundations of Public Policy*. Nueva York: Eldar Shafir, 2012.

TREVIÑO, LINDA, GARY WEAVER Y SCOTT REYNOLDS. "Behavioral Ethics in Organizations: A Review". *Journal of Management*, vol. 32, n.º 6, 2006.

UNIÓN EUROPEA. *Behavioural Insights Applied to Policy*. Luxemburgo: Publications Office of the European Union, 2016.

UNODC. *State of Integrity: A guide on conducting corruption risk assessments in public organizations*. Viena: Naciones Unidas, 2009.

VAN MONTFORT, ANDRÉ, LAURA BECK Y ANNEKE TWINJSTRA. "Can integrity be taught in public organizations?". *Public Integrity*, vol. 15, n.º 2, 2013.

WEINSCHENK, SUSAN. *Diseño inteligente. Las 100 cosas que todo diseñador necesita saber sobre las personas*. Madrid: Anaya Multimedia, 2021.

YEUNG, KAREN. "'Hypernudge': Big Data as a mode of regulation by design". *Information, Communication & Society*, vol. 20, n.º 1, 2017.

ZIMMERMANN, SARAH. "¿Incentivo o manipulación?". *Mente y Cerebro*, n.º 78, 2016.