

# Servicios de Inteligencia Artificial orientados a los documentos

Fairness, accountability and transparency en la obtención de datos

Fecha 22/04/2024

# ÍNDICE

- 01 Registro de miembros del equipo.
- 02 Registro de las entregas de algoritmos.
- 03 Registro de los datos empleados para las entregas.
- 04 Registro de aplicaciones integradas.
- 05 Registro de sesgos identificados.
- 06 Categorías de los sistemas de información.

A continuación se detallan abreviaturas que se utilizarán a lo largo del documento

- ANON: servicio de Anonimizado de documentos
- CD: servicio de Clasificación Documental
- DIG: procesos de digitalización
- NER: servicio de Identificación de Entidades Nominales
- RE: servicio de Relación entre Entidades Nominales
- SGP: sistema de gestión procesal
- SIM: similaridad de personas físicas o jurídicas
- RD-CDCJ: Resúmenes de documentos de Cuentas de Consignaciones Judiciales
- CR: servicio de Clasificación de Clases Registrales
- CO: servicio de Clasificación de Órdenes Jurisdiccionales
- FR: servicio de Fechas Relevantes
- CE: servicio de Cuantías económicas destacadas.
- SI: servicio de síntesis de documentos.
- OS: orquestador de servicios.
- LF: lectura fácil.

Continúa

- TR: Servicio de traducciones

# REGISTRO DE MIEMBROS DEL EQUIPO - I

El Registro de miembros del equipo contiene datos personales e identificadores de red, que no se incluyen aquí por razones de protección de datos.

Iniciales	Sexo	Fecha de ingreso	Fecha de baja	Rol	Áreas de intervención	Fecha de registro
AMPT	H	29/01/2018		Director	CD, NER, ANON, RE, SIM, CR, RD-CDCJ, DIG, CR, CO, FR, CE, SI,TR	23/11/2021
DDL	H	29/01/2018	01/07/2022	Director técnico	CD, NER, ANON	23/11/2021
MJR	H	30/07/2018	22/09/2021	Consultor Big-Data	CD, NER, ANON	23/11/2021
DGM	H	02/09/2019	31/08/2020	Desarrollador IA & NLP	CD, NER, ANON	23/11/2021
AMLS	H	18/05/2020	01/04/2022	Arquitecto IA & NLP	CD, NER, ANON	23/11/2021
TAF	M	28/01/2021		Jurista-Documentalista	CD, CR	23/11/2021
CJP	M	28/01/2021		Jurista-Documentalista	CD, CR	23/11/2021
MDMC	M	15/02/2021	09/10/2023	Jurista-Documentalista	CD, CR	23/11/2021
SMR	M	09/03/2021	01/06/2022	Admvo-Documentalista	NER	23/11/2021
VRM	M	09/03/2021	07/11/2021	Admvo-Documentalista	NER	23/11/2021
MTDG	M	09/03/2021	01/06/2022	Admvo-Documentalista	NER	23/11/2021
CCN	M	09/03/2021	01/06/2022	Admvo-Documentalista	NER	23/11/2021
SJG	M	09/03/2021	01/06/2022	Admvo-Documentalista	NER	23/11/2021
CRC	H	08/04/2021		Analista servicios Python	CD, NER, ANON, RE, SIM, CR, CO, FR, CE, SI,TR	23/11/2021
LAGA	H	21/06/2021		Consultor IA & Datos	CD, NER, ANON, SIM, CR, FR, CE, SI, TR	23/11/2021
OGG	H	23/11/2021	07/02/2022	Arquitecto DevOPS	CD, NER, ANON	25/11/2021
FCP	H	07/03/2022	30/06/2022	Analista Programador	CD, NER	08/03/2022
ASR	M	04/04/2022	31/12/2022	Consultor Data Science	DIG	13/04/2022

# REGISTRO DE MIEMBROS DEL EQUIPO - II

El Registro de miembros del equipo contiene datos personales e identificadores de red, que no se incluyen aquí por razones de protección de datos.

Iniciales	Sexo	Fecha de ingreso	Fecha de baja	Rol	Áreas de intervención	Fecha de registro
MEP	M	08/04/2022		Jefe de Proyecto	CD, NER, ANON	13/06/2022
FMM	H	08/04/2022		Arquitecto Kubernetes	CD, NER, ANON, SIM, CR, FR, CE, SI, OS, TR	13/06/2022
OTM	H	01/06/2022	01/02/2023	Jefe de Proyecto	CD, NER, ANON	13/06/2022
CAA	H	16/06/2022	15/05/2023	Jefe de Proyecto	CD, RE, SIM	08/07/2022
RLC	H	20/06/2022	28/04/2023	Analista Programador	CD, RE, SIM	08/07/2022
SSG	H	20/06/2022		Programador Python	CD, RE, SIM, FR, CE	08/07/2022
JZ	H	12/07/2022	14/11/2023	Analista Programador	CD, RE, SIM, CR, FR, CE, SI, ANON	08/07/2022
PQC	H	12/07/2022	28/04/2023	Data Sciencs	CD, RE, TR	08/07/2022
FDD	H	12/07/2022	28/04/2023	Data Sciencs	CD, RE	08/07/2022
BAV	H	29/07/2022	31/12/2022	Jurista	RD-CDCJ	06/10/2022
MGN	M	29/07/2022	31/12/2022	Jefe de Proyecto	RD-CDCJ	06/10/2022
NAG	M	29/07/2022	31/12/2022	Admvo-Documentalista	RD-CDCJ	06/10/2022
CDC	H	29/07/2022	31/12/2022	Analista Programador	RD-CDCJ	06/10/2022
JMOG	H	11/08/2022	31/12/2022	Analista Programador	RD-CDCJ	06/10/2022
ABJ	H	21/09/2022	31/12/2022	Data Sciencs	RD-CDCJ	06/10/2022
LAR	M	21/09/2022	28/04/2023	Data Sciencs	RE, TR	06/10/2022
AVS	H	21/09/2022	31/12/2022	Data Sciencs	RD-CDCJ	06/10/2022
NLF	M	26/09/2022	31/12/2022	Admvo-Documentalista	RD-CDCJ	06/10/2022

# REGISTRO DE MIEMBROS DEL EQUIPO - III

El Registro de miembros del equipo contiene datos personales e identificadores de red, que no se incluyen aquí por razones de protección de datos.

Iniciales	Sexo	Fecha de ingreso	Fecha de baja	Rol	Áreas de intervención	Fecha de registro
LRP	M	26/09/2022	31/05/2022	Admvo-Documentalista	RD-CDCJ	06/10/2022
MDEA	M	03/10/2022		Jurista	CR, TR	06/10/2022
RMAR	M	25/10/2022		Jurista	CR	04/11/2022
MSG	M	04/11/2022	11/01/2023	Jurista-Documentalista	RD-CDCJ	04/11/2022
MEHG	M	30/01/2023		Jurista	CR	06/02/2023
ESC	H	05/06/2023		Analista Programador	OS	14/06/2023
PMP	M	18/07/2023		Educadora Social	LF	26/07/2023
ARH	H	01/08/2023	12/02/2024	Data Scients	RE, SI, TR	02/08/2023
RGL	H	08/01/2024		Programador Front-End	ANOM	09/02/2024
ISL	H	20/03/2024		Data Scients	RE, SI, TR	01/04/2024

## TECNOLOGÍAS COMUNES UTILIZADAS EN EL DESARROLLO DE ALGORITMOS

**Python:** es un lenguaje de programación interpretado cuya filosofía hace hincapié en la legibilidad de su código. Se trata de un lenguaje de programación multiparadigma, ya que soporta parcialmente la orientación a objetos, programación imperativa y, en menor medida, programación funcional. Es un lenguaje interpretado, dinámico y multiplataforma.

**Flask:** Es un “micro” Framework escrito en Python y desarrollado para simplificar y hacer más fácil la creación de Aplicaciones Web bajo el patrón MVC.

**Requests:** Es una sencilla biblioteca HTTP para Python.

**Gunicorn:** Es un servidor HTTP WSGI de Python para UNIX. Es un modelo de trabajador pre-fork. El servidor Gunicorn es ampliamente compatible con varios frameworks web, se implementa de forma sencilla, consume pocos recursos del servidor y es bastante rápido.

**Servicios coordinadores:** conjunto de servicios que se encargan de orquestar y coordinar el resto de servicios específicos. Implementan servicios rest basados en Flask, con llamadas a los servicios mediante la biblioteca Requets.

## ESPECÍFICAS PARA NER & ANON

**Pymupdf:** Es un enlace de Python para MuPDF, un visor, renderizador y kit de herramientas ligero para PDF, XPS y libros electrónicos, mantenido y desarrollado por Artifex Software, Inc.

**Shapely:** Es un paquete de Python con licencia BSD para la manipulación y el análisis de objetos geométricos planos. Se basa en las bibliotecas GEOS (el motor de PostGIS) y JTS (de la que se ha portado GEOS), ampliamente extendidas.

**Ocrmypdf:** Biblioteca de Python que utiliza el motor de OCR de Tesseract, genera un archivo PDF/A con capacidad de búsqueda a partir de un PDF normal que sólo contenga imágenes, permitiendo su búsqueda.

**Tesseract:** es un motor de reconocimiento óptico de caracteres para varios sistemas operativos, su desarrollo es financiado por Google desde el 2006.

**Docx2python:** Biblioteca para la manipulación de documentos docx, que permite extraer y editar cabeceras, pies de página, texto, notas a pie de página, notas finales, propiedades e imágenes.

**Pyth3:** pretende facilitar la manipulación de texto marcado de diferentes formatos comunes, tales como RTF, XHTML y texto plano,

## ESPECÍFICAS PARA CD

**Torch:** Un marco de aprendizaje automático de código abierto que acelera el camino desde la creación de prototipos de investigación hasta el despliegue de producción.

**Transformers:** Transformers proporciona miles de modelos preentrenados para realizar tareas en diferentes modalidades como texto, visión y audio, mantenido y desarrollado por Hugging Face.

## ESPECÍFICAS PARA EXTRACCIÓN DE NER

**Torch:** Un marco de aprendizaje automático de código abierto que acelera el camino desde la creación de prototipos de investigación hasta el despliegue de producción.

**Transformers:** proporciona miles de modelos preentrenados para realizar tareas en diferentes modalidades como texto, visión y audio, mantenido y desarrollado por Hugging Face.

**Tokenizers:** proporciona una implementación de los tokenizadores más utilizados en la actualidad, con un enfoque en el rendimiento y la versatilidad, es un enlace sobre la implementación de Rust , mantenido y desarrollado por Hugging Face.

# REGISTRO DE LAS ENTREGAS DE ALGORITMOS I

El Registro de algoritmos entregados puede seguirse en la siguiente tabla.

SERVICIO	DESCRIPCIÓN	VERSIÓN	ENTREGA	ALGORITMO & HASH	NOVEDADES	REGISTRO
Clasificador documental	Servicios que permite determinar la familia y subfamilia del mapa documental al que pertenece un documento dado.	Prototipo	Marzo-22	Sha-256	Primera versión que contempla 71 subtipos documentales	08/03/2022
Extractor de entidades nominales	Permitirá detectar las distintas categorías de entidades nominales identificadas dentro de los documentos judiciales	Prototipo	Marzo-22	Sha-256	Primera versión que identificará: personas, roles, organizaciones, lugares, fechas, referencias legales, cantidades, números de identificación, códigos	08/03/2022
Clasificador documental	Servicios que permiten determinar la familia y subfamilia del mapa documental. Documentos de Abogados, Procuradores y Graduados Sociales	1.0	15-Junio-22	N/A	Versión productiva liberada con los 118 subtipos documentales del juego.	26/07/2022
Extractor de entidades nominales	Permitirá detectar las distintas categorías de entidades nominales identificadas dentro de los documentos judiciales	1.0	15-Junio-22	N/A	Versión productiva liberada con la incorporación de los últimos etiquetados	26/07/2022
Clasificador Documental	Mejoras en la clasificación documental, con optimización del árbol inicial de clasificado de documentos de Abogados, Procuradores y Graduados Sociales	1.1	29/08/2022	N/A	Se simplifica el modelo desplegado en la 1.0 para un menor coste de memoria.	25/01/2023
Extractor de entidades nominales	Mejoras en la extracción del taggeado de entidades para la identificación.	1.1	29/08/2022	N/A	Se corrigen deficiencias observadas en el paso por el tokenizador de elementos	25/01/2023

# REGISTRO DE LAS ENTREGAS DE ALGORITMOS II

El Registro de algoritmos entregados puede seguirse en la siguiente tabla.

SERVICIO	DESCRIPCIÓN	VERSIÓN	ENTREGA	ALGORITMO & HASH	NOVEDADES	REGISTRO
Clasificador documental	Se amplían los servicios de clasificado a documentos de hospitales, FyCSE, Instituciones Penitenciarias y de Menores, Fiscalías en un grupo y los Judiciales en otro grupo distinto.	1.2	30/12/2022	<p>Grupo 1 Sha1sum: 4a7e52e165c1dbcbcc89851742df57df87f529c9a5197d1add57baf2fff9d2dba389ded08b3eb1c5370e2917723107ebee6c2632380e2cb767e81671088a6fb</p> <p>Grupo 2 Sha1sum: 9b6149a21961d5064b53585abda537ddc1747ba617a92f04a8aa5963041fda5e57d86865652523173bb9931bdf481bbeb2662d3f8f7c2eccc1eece6f3f7784b6</p> <p>Grupo 3 Sha1sum: 8dc0c19123854ce8b1d53749efd3df4bf370ff8b46e38211c194e0e301bce47371e992dfbcfc395743f71e042e5225306898a1455b9e162e4ecc78f708b4b5ac</p>	Ampliación de grupos de clases documentales a clasificar.	27/02/2023
Similaridad	Se implantan procesos de similaridad para identificar parecidos entre entidades nominales de personas físicas y jurídicas identificadas	1.0	30/12/2022	N/A	Identificar diferentes nomenclaturas para la llamada de las personas físicas o jurídicas.	27/02/2023
Fechas Relevantes	Servicio que permite extraer las fechas destacadas y normalizadas de un documento ordenadas de modo cronológico.	1.0	03/05/2023	N/A		03/05/2023

El Registro de algoritmos entregados puede seguirse en la siguiente tabla.

SERVICIO	DESCRIPCIÓN	VERSIÓN	ENTREGA	ALGORITMO & HASH	NOVEDADES	REGISTRO
Extractor de entidades nominales	Permitirá detectar las distintas categorías de entidades nominales identificadas dentro de los documentos judiciales	1.2	03/05/2023	Sha1sum: dbaa0da1ecb6020332d6aa8d42b541577039db6e2ab4efea8df52881656bb0ad1cc09664344b4063c27b1d38df19d52eb56f2d070ca13e6bf7155e4ad147a7e2 Direcciones Sha1sum: 1d05829489815480a87670b14254fca5cf1d95b479b852d7e9b71e7c94b95c26641650192e591c70250b219c8aa546154b6f1fd0ae6f1e03df1e356453c7c2a8	Corrección de funciones de identificación y reemplazo de entidades en generación de documentos anonimizados espejo.	27/02/2023
Clasificador de Órdenes Jurisdiccionales	Servicio que permite clasificar un documento por el Orden Jurisdiccional de destino.	1.0	03/05/2023	Sha1sum: 0f63764def5edc13629aa25e9472c6c17de6ffedcb84be07ea157558f3ea4e5f56191ed8b25dc82ae04399e144d815200b7a32dd3404d509035dc65c0e1c75ab		03/05/2023
Cuantías destacadas	Servicio que permite extraer las cuantías económicas destacadas de un documento.	1.0	03/05/2023	N/A		03/05/2023

# REGISTRO DE LAS ENTREGAS DE ALGORITMOS IV

El Registro de algoritmos entregados puede seguirse en la siguiente tabla.

SERVICIO	DESCRIPCIÓN	VERSIÓN	ENTREGA	ALGORITMO & HASH	NOVEDADES	REGISTRO
Imputador de clases de Registro	Servicio que permite imputar la clase de Registro de un documento iniciador. Órdenes Civil y Social.	1.0	03/05/2023	Civil Sha1sum: ffaa00558e55a2116faaf590884b4b 57687440958d75b1b494f9d370ad6 5d78d2f04d161b7d5183d3ae0ae96 d7a72c2ddf8fa116db59a9c600691a c5d6e5b092 Social Sha1sum: c00410441f2bd725d446d65154365 56caf407c38c6e20caa3ff2f1c539bd 16ee7d31a709a51ce5225b2fd8eca d144815ce51bd5271b6e3a5caae32 c83891c631		03/05/2023
Relación de entidades.	Servicio que permite establecer las relaciones entre las entidades nominales descubiertas. Las relaciones se establecen en dos niveles: el nivel 1 trata de hallar todas las propiedades que pueden ser atribuidas a un objeto persona; mientras que el nivel 2 establece las relaciones de las personas entre sí. Esto por ejemplo ayudará en el descubrimiento de los intervinientes en un procedimiento.	1.0	03/05/2023	Sha1sum: 742ea6cf6dbd25488f04d6a9896495 602b371714d72d6f06f3f19f4d62e6 a21a3c2e513356b06c4393c31b4fe4 6fd228f921723f21a6eab03ebee5bf 1c6b3ab7		03/05/2023
Síntesis de documentos.	Se realiza un primer prototipo de servicio para la síntesis de documentos empleando modelos de lenguaje grandes (LLM), en dos versiones: resumen y lenguaje sencillo.	1.0	03/05/2023	N/A		03/05/2023

# REGISTRO DE LAS ENTREGAS DE ALGORITMOS V

El Registro de algoritmos entregados puede seguirse en la siguiente tabla.

SERVICIO	DESCRIPCIÓN	VERSIÓN	ENTREGA	ALGORITMO & HASH	NOVEDADES	REGISTRO
Traductor	Servicio de traducción de documentos a lenguas cooficiales. Se presta soporte a la traducción bidireccional a/desde: Catalán, Euskera, Gallego y Valenciano.	1.0	19/10/2023	<p>CAS-CAT Sha1sum:                      7507988f2f3b7004ab57415fdb8144                      65ae4c41c35fea8c5e0054c154bc57                      ad1a710f224ba1d68922f4421c2909                      6df04f840672c77e61ba8f9181994d                      e92f02db</p> <p>CAT-CAS Sha1sum:                      a2cd5421e9e5e645cacd6f2ba35625                      5820baa1ca1e1b6c6529bf40f3cc5fd                      157d3bdd8c41eb1d176c1ccda137b                      9d85b1ab6a208056113f2f731e373                      00f48572c</p> <p>CAS-EUS Sha1sum:                      758f4781da5037f69d084c2f60391b                      bed5b7ebf4951c20a6b8e1c260f6b2                      a08ff74fe3264f9382a113644f66331                      7597eb6ea2d675262bc6ab8d14164                      f6df9809</p> <p>EUS-CAS Sha1sum:                      6e0fb025731a55928902fbda1d967                      a51d2cbc0cfa53cd3575bcd1811d1                      3ab6834ff6822e124d2776c7957e17                      e447a2fed932d76f4e9225284e36a6                      6a10f812f</p>		23/10/2023

# REGISTRO DE LAS ENTREGAS DE ALGORITMOS VI

El Registro de algoritmos entregados puede seguirse en la siguiente tabla.

SERVICIO	DESCRIPCIÓN	VERSIÓN	ENTREGA	ALGORITMO & HASH	NOVEDADES	REGISTRO
Traductor	Continúa			CAS-GAL Sha1sum: 8320092cef3baad0351b8e5303f49c 103f9932a2f1d1c37566b9fb630b92 3b501b6149332d09aa2745fe496fb a151c1867e24bf408d365358b75aaf 116ea0b32 GAL-CAS Sha1sum: 7f03996a55707ba7ad7a0e546c77bf 0f7963e32547616131a147a747dd9 3b8bd9f04bb5e90b997b73e9409db 732757660fc642c68f7c255e58abb8 059414e80d CAS-VAL Sha1sum: 7507988f2f3b7004ab57415fdb8144 65ae4c41c35fea8c5e0054c154bc57 ad1a710f224ba1d68922f4421c2909 6df04f840672c77e61ba8f9181994d e92f02db VAL-CAS Sha1sum: a2cd5421e9e5e645cacd6f2ba35625 5820baa1ca1e1b6c6529bf40f3cc5fd 157d3bdd8c41eb1d176c1ccda137b 9d85b1ab6a208056113f2f731e373 00f48572c		

El Registro de algoritmos entregados puede seguirse en la siguiente tabla.

SERVICIO	DESCRIPCIÓN	VERSIÓN	ENTREGA	ALGORITMO & HASH	NOVEDADES	REGISTRO
Extractor de entidades nominales	Permitirá detectar las distintas categorías de entidades nominales identificadas dentro de los documentos judiciales	1.3	19/10/2023	N/A	Se realizan mejoras en los sistemas de procesado de documentos, con la posibilidad de eliminar las imágenes en el procesado de los mismos.	23/10/2023
Imputador de clases de Registro	Servicio que permite imputar la clase de Registro de un documento iniciador. Órdenes Civil, Social y Mercantil.	1.1	19/10/2023	Mercantil Sha1sum: 544e089a244abfa52e75440def093190c73cb1fd8e94a231c86c7da562244bc38a8c36d52f096c29abd2fba39e147183ae0a0389d4b850351982dff54dabad0	Se incorporan las clases de registro de tipología Mercantil.	23/10/2023
Relación de entidades	Servicio que permite establecer las relaciones entre las entidades nominales descubiertas. Las relaciones se establecen en dos niveles: el nivel 1 trata de hallar todas las propiedades que pueden ser atribuidas a un objeto persona; mientras que el nivel 2 establece las relaciones de las personas entre sí. Esto por ejemplo ayudará en el descubrimiento de los intervinientes en un procedimiento.	1.1	19/10/2023	Direcciones Sha1sum: 1d05829489815480a87670b14254fca5cf1d95b479b852d7e9b71e7c94b95c26641650192e591c70250b219c8aa546154b6f1fd0ae6f1e03df1e356453c7c2a8	Se han modificado y añadido determinados elementos de las respuestas del servicio, para normalizado de direcciones y de los formateos de nombre y apellidos.	23/10/2023

El Registro de algoritmos entregados puede seguirse en la siguiente tabla.

SERVICIO	DESCRIPCIÓN	VERSIÓN	ENTREGA	ALGORITMO & HASH	NOVEDADES	REGISTRO
Síntesis de documentos	Superada la primera fase del prototipo se despliegan versiones ya entregables. Unas se han realizado con sistemas de ranqueo, otras con entrenamientos de IA y otras mediante uso de LLM.	1.1	19/10/2023	<p>Extracto Sha1sum:                      693b8682759bdf7db7c9abb46f79df                      8c30131d0ca8309374e3976b5d72e                      2019857d6fd1ab6781bc3eb83d096                      b8a48c780359a2d4e307075bb9871                      85c0f5829c8</p> <p>Sencillo Sha1sum:                      9404aabc22c33332f301fafd66f32ea                      81e806a8326ac6e2cc5eaf434b383f                      b600915b85dfff28911d2c32b47069                      7178a4795f7b1af110c2885aed1f9e                      b151e5c</p>	Se despliegan el resumen de extracto (titulares), el de sentencias y autos, el de documentos genéricos, así como el de documentos sencillos.	23/10/2023

# REGISTRO DE LOS DATOS EMPLEADOS PARA LAS ENTREGAS I

La elección de documentos para su tratamiento se realizó de forma que quedaran equilibradas las siguientes variables:

- Cantidad de documentos por cada provincia (partido judicial) seleccionada.
- Cantidad de documentos por orden jurisdiccional.
- Cantidad de documentos por año.

La cifra que se refleja finalmente en el registro es el monto total de documentos que han pasado a formar parte del corpus de entrenamiento de la versión, habiendo sido la cifra de partida mucho más amplia.

SERVICIO	VERSIÓN	NÚMERO DE DOCUMENTOS	ORIGEN	ÁMBITO	INTERVALO TEMPORAL	CRITERIOS FUNCIONALES	CRITERIOS TÉCNICOS	REGISTRO
CD	Prototipo	16713	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Se excluyen del dataset las siguientes tipologías: <ul style="list-style-type: none"> <li>• Históricos de Minerva</li> <li>• Diligencias de reparto</li> <li>• Diligencias de devolución</li> <li>• Documentación en papel</li> <li>• Documentación no digitalizable</li> <li>• Documentación indeterminada</li> <li>• Subtipos documentales exclusivos de la jurisdicción militar.</li> </ul>	Se excluyeron del dataset: <ul style="list-style-type: none"> <li>- Documentos con ruta o clasificación desconocida</li> <li>- Documentos con más de una clasificación</li> <li>- Subtipos no vigentes del mapa documental</li> <li>- Documentos cuya extensión no eran pdf</li> </ul>	08/03/2022
NER	Prototipo	6647	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Descartados subtipos de carácter judicial y subtipos con estado no vigente. Documentos con contenido en la mayor parte de su texto con estructura jurídica.	Selección exclusiva de formatos PDF no escaneados, o con OCR previo.	08/03/2022

# REGISTRO DE LOS DATOS EMPLEADOS PARA LAS ENTREGAS II

SERVICIO	VERSIÓN	NÚMERO DE DOCUMENTOS	ORIGEN	ÁMBITO	INTERVALO TEMPORAL	CRITERIOS FUNCIONALES	CRITERIOS TÉCNICOS	REGISTRO
CD	1.0	14601	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Solo se incluyen en el dataset clases documentales pertenecientes a los colectivos: Abogados, Procuradores y Graduados Sociales.	Mismos criterios que en prototipo	26/07/2022
NER	1.0	6647	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Mismos criterios que en prototipo	Mismos criterios que en prototipo	26/07/2022
CD	1.1	14601	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Solo se incluyen en el dataset clases documentales pertenecientes a los colectivos: Abogados, Procuradores y Graduados Sociales.	Mismos criterios que en prototipo	25/01/2023
NER	1.1	6647	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Mismos criterios que en prototipo	Mismos criterios que en prototipo	25/01/2023
CD	1.2	2292 para el grupo de hospitales, FyCSE,... 10260 Judiciales	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Se incluyen clases documentales de: hospitales, FyCSE, Instituciones Penitenciarias y de Menores y Fiscalías. También se construyen clases documentales Judiciales.	Mismos criterios que en prototipo	27/02/2023
NER	1.2	6647	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Mismos criterios que en prototipo	Mismos criterios que en prototipo	27/02/2023
SIM	1.0	6647	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Descartados subtipos de carácter judicial y subtipos con estado no vigente. Documentos con contenido en la mayor parte de su texto con estructura jurídica.	Selección exclusiva de formatos PDF no escaneados, o con OCR previo.	27/02/2023

# REGISTRO DE LOS DATOS EMPLEADOS PARA LAS ENTREGAS III

SERVICIO	VERSIÓN	NÚMERO DE DOCUMENTOS	ORIGEN	ÁMBITO	INTERVALO TEMPORAL	CRITERIOS FUNCIONALES	CRITERIOS TÉCNICOS	REGISTRO
CO	1.0	27000	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Documentos no judiciales de todas las tipologías.	Se excluyeron del dataset: <ul style="list-style-type: none"> <li>- Documentos con ruta o clasificación desconocida</li> <li>- Documentos con más de una clasificación</li> <li>- Subtipos no vigentes del mapa documental</li> <li>- Documentos cuya extensión no eran pdf</li> </ul>	03/05/2023
CR	1.0	14000 Civil + 9000 Social	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Documentos iniciadores con destino al orden Civil (no mercantil) y Social.	Se excluyeron del dataset: <ul style="list-style-type: none"> <li>- Documentos con ruta o clasificación desconocida</li> <li>- Documentos con más de una clasificación</li> <li>- Subtipos no vigentes del mapa documental</li> <li>- Documentos cuya extensión no eran pdf</li> </ul>	03/05/2023

# REGISTRO DE LOS DATOS EMPLEADOS PARA LAS ENTREGAS III

SERVICIO	VERSIÓN	NÚMERO DE DOCUMENTOS	ORIGEN	ÁMBITO	INTERVALO TEMPORAL	CRITERIOS FUNCIONALES	CRITERIOS TÉCNICOS	REGISTRO
RE	1.0	4000	SGP Minerva	Murcia, Toledo y Valladolid	2010-2020	Documentos identificados en NER, con variedad y dispersión de las entidades en ellos contenidas.	Se excluyeron del dataset: <ul style="list-style-type: none"> <li>- Documentos con ruta o clasificación desconocida</li> <li>- Documentos con más de una clasificación</li> <li>- Subtipos no vigentes del mapa documental</li> <li>- Documentos cuya extensión no eran pdf</li> </ul>	03/05/2023
SI	1.1	68886	CENDOJ	Tribunal Supremo, Audiencia Nacional y Tribunales Superiores de Justicia	Anterior a 2023	Documentos anonimizados descargados del CENDOJ que tienen resúmenes de titulares realizados	Se han escogido documentos de modo aleatorios de los distintos órdenes jurisdiccionales, tanto autos como sentencias.	23/10/2023

# REGISTRO DE APLICACIONES O SERVICIOS INTEGRADOS I

Registro de aquellas aplicaciones o servicios que van a hacer uso de algoritmos de Inteligencia Artificial originados por entrenamiento de datasets generados internamente.

SERVICIO	APLICACIÓN O SERVICIO	FINALIDAD	INICIO / FIN	DATOS PERSONALES	CATEGORIAS	APLICA GDPR	REGISTRO
NER	ANON	Extraer datos personales de los documentos con la finalidad de reemplazarlos/sustituirlos/ocultarlos a fin de salvaguardar a las personas de su difusión o publicidad.	ENE-22/FEB-23	SI	Nombres y apellidos Domicilios Identificadores de documentos personales Identificadores de internet (mails) Identificadores de vehículos Códigos seguros de verificación de documentos donde figuren.	SI	08/03/2022
CD	LexNET	LexNET utilizará el clasificador documental para obtener la clase de los documentos que va a procesar	ABR-22/FEB-23	NO	No se harán uso de los datos personales	NO	08/03/2022
ANON	SIM	Identificar personas físicas o jurídicas parecidas para realizar anonimizados similares en todas.	SEP-22/FEB-23	SI	Nombres y apellidos	SI	27/02/2023
NER	FR	Identificar las fechas relevantes que aparecen en un documento ordenadas de modo cronológico	ENE-23/MAYO-23	NO	No se hacen uso de datos personales en el tratamiento.	NO	03/05/2023
NER	FR	Identificar las cuantías económicas destacadas de un documento para poder relacionarlas.	ENE-23/MAYO-23	NO	No se hacen uso de datos personales en el tratamiento.	NO	03/05/2023

# REGISTRO DE APLICACIONES O SERVICIOS INTEGRADOS II

Registro de aquellas aplicaciones o servicios que van a hacer uso de algoritmos de Inteligencia Artificial originados por entrenamiento de datasets generados internamente.

SERVICIO	APLICACIÓN O SERVICIO	FINALIDAD	INICIO / FIN	DATOS PERSONALES	CATEGORIAS	APLICA GDPR	REGISTRO
NER	RE	Relacionar las entidades localizadas entre sí, como propiedades de las personas, o de éstas entre sí.	ENE-23/MAYO-23	SI	Nombres y apellidos Domicilios Identificadores de documentos personales Identificadores de internet (mails) Identificadores de vehículos Códigos seguros de verificación de documentos donde figuren.	SI	03/05/2023
ANON	SI	Obtener de un documento anonimizado una síntesis del mismo en dos versiones: resumen y sencilla, basándose en el uso de librerías de modelos de lenguaje grandes.	ENE-23/MAYO-23	SI	Nombres y apellidos Domicilios Identificadores de documentos personales Identificadores de internet (mails) Identificadores de vehículos Códigos seguros de verificación de documentos donde figuren.	SI	03/05/2023
NER, RE	LexNET	Localización de la información de los intervinientes y sus relaciones.	SEP-23/NOV-23	SI	Nombres y apellidos Domicilios Identificadores de documentos personales	SI	23/10/2023
CR	Robotización (RPA)	Obtención de la clase de registro de iniciadores de procedimientos Monitorios	SEP-23/OCT-23	NO		NO	23/10/2023

# REGISTRO DE APLICACIONES O SERVICIOS INTEGRADOS III

Registro de aquellas aplicaciones o servicios que van a hacer uso de algoritmos de Inteligencia Artificial originados por entrenamiento de datasets generados internamente.

SERVICIO	APLICACIÓN O SERVICIO	FINALIDAD	INICIO / FIN	DATOS PERSONALES	CATEGORIAS	APLICA GDPR	REGISTRO
SI	Carpeta Justicia	Obtener diferentes clases de resúmenes de documentos.	OCT-23/DIC-23	SI	Nombres y apellidos Domicilios Identificadores de documentos personales Identificadores de internet (mails) Identificadores de vehículos Códigos seguros de verificación de documentos donde figuren.	SI	23/10/2023
ANON	Carpeta Justicia	Obtener documentos anonimizados.	SEP-23/DIC-23	SI	Nombres y apellidos Domicilios Identificadores de documentos personales Identificadores de internet (mails) Identificadores de vehículos Códigos seguros de verificación de documentos donde figuren.	SI	23/10/2023

El registro de sesgos identifica si para la fabricación, diseño o composición de los servicios de IA se ha utilizado o permitido algún sesgo de los contenidos en los documentos empleados. Asimismo, se detallarán los sesgos de carácter técnico que se han tenido en cuenta.

Referencias legales que precisan de la identificación de sesgos en los documentos:

- Existen distintos preceptos en la **Constitución Española** que abogan por la **igualdad de las personas** en todos los ámbitos de su vida, incluido en el ámbito jurídico, y ante la ley. Dichos preceptos se recogen en los artículos 9.2, 14 y 24.2 de la Constitución Española.
- De manera amplia los **datos personales de todas las personas están protegidos para su tratamiento** por la Ley Orgánica 3/2018 de Protección de datos personales y garantía de los derechos digitales que, además, adapta la normativa española al Reglamento de la Unión Europea 2016/679<sup>1</sup> relativo a la protección de las personas físicas en lo que respecta al tratamiento de los datos personales y a la libre circulación de estos datos<sup>[1]</sup>.

En concreto, existen determinados datos personales que están **sometidos a una especial protección**, estos datos tienen como finalidad principal identificar su ideología, afiliación sindical, religión, orientación sexual, creencias u origen racial o étnico (art.9.1 LO 3/2018).

- La **protección de los menores de edad** se ve reflejada en otras normas como Ley Orgánica 8/2021, de 4 de junio, de protección integral a la infancia y la adolescencia frente a la violencia (que no incluye transexuales o niños que no se identifiquen con su sexo de nacimiento.) Su principal objetivo es garantizar los derechos fundamentales de los niños, niñas y adolescentes sin tener en cuenta su color de piel y origen social, frente a cualquier tipo de violencia. Todo ello, través de una serie de medidas de protección integral que incluyen desde la detección precoz hasta la reparación de los daños.

Por otro lado, en lo referido a la tramitación de los procedimientos judiciales, y en consecuencia al tratamiento de los documentos y datos obrantes en ellos, **existen casos en los que es preciso conocer en detalle determinados aspectos de la vida de las personas, aunque sean datos sensibles especialmente protegidos**. A continuación, se muestran algunos datos imprescindibles de ser tenidos en cuenta en determinados procedimientos judiciales, sin ser un *“numerus clausus”*

- Referencias a la **edad** (art.19, 69 o 140.1, entre otros, del Código Penal, art. 1 Ley Orgánica de responsabilidad penal del menor, art.16.2 Ley jurisdicción social, art.770.4ª de la Ley de Enjuiciamiento Civil, derecho a ser oído en juicio, derecho a declarar art. 152CC...). Por ejemplo, para la exigencia de la responsabilidad penal de la persona ante la comisión de un hecho delictivo en función de la edad del autor. En caso de que el autor sea menor de edad, dicha responsabilidad se regula con la [Ley Orgánica de Responsabilidad Penal del Menor](#) 5/2000 de 12 de enero, en caso de que el autor sea mayor de edad regirá lo dispuesto en el Código Penal.
- Referencias a la **capacidad modificada judicialmente de las personas** (art. 140.1 156 bis 4.b, 156 ter, 165 172 o 187, entre otros, del Código Penal. art.7.2 de la Ley de Enjuiciamiento Civil, Ley 8/2021, de 2 de junio, por la que se reforma la legislación civil y procesal para el apoyo a las personas con discapacidad en el ejercicio de su capacidad jurídica.)
- Referencias **cargos o profesiones** (art. 142.1 157, 175 o 199 entre otros del Código Penal). Por ejemplo, para el enjuiciamiento de determinadas actividades laborales en el orden social.
- Referencias a **raza, nacionalidad, orientación sexual, religión o ideología, lenguaje, localización de residencia** (art. 170, 314 o 510, entre otros, del Código Penal). En la comisión de delitos de odio (infracción o acto penal motivado por prejuicios contra una o varias personas por el hecho de pertenecer a un determinado grupo social (raza, lenguaje, color, religión, orientación sexual o similar.) es necesario conocer determinadas características atribuidas a las personas para la correcta calificación de los hechos. Estos delitos tienen como objeto la protección de los derechos fundamentales y de las libertades de determinados colectivos en situación de discriminación y desigualdad, garantizados por la Constitución. En esta línea, se han promulgado distintas leyes que discriminan a favor del colectivo vulnerable, como la Ley 19/2020, de 30 de diciembre, de igualdad de trato y no discriminación, lo que supone un sesgo en sí misma.

Además de las normas dictaminadas, pueden existir sesgos o referencias en los procedimientos judiciales que identifican a un colectivo de personas, que en algunos casos puede asociarse a una mayor vulnerabilidad a dichos colectivos lo que se traduce en una mayor sensibilización de la sociedad para dotar a estos colectivos de mayor protección.

- Referencias al **sexo femenino** de las personas (art. 144,148, 145 161, 172.2,153, 148, 510, 314, entre otros, del Código Penal). En los delitos de violencia de género es imprescindible tomar en cuenta el sexo de la víctima, así la Ley Orgánica 1/2004, de 28 de diciembre, de Medidas de Protección Integral contra la Violencia de Género) describe estos hechos como “[...]manifestación de la discriminación, la situación de desigualdad y las relaciones de poder de los **hombres sobre las mujeres**, se ejerce sobre éstas por parte de quienes sean o hayan sido sus cónyuges o de quienes estén o hayan estado ligados a ellas por relaciones similares de afectividad, aun sin convivencia.”
- Referencias a las **relaciones de afectividad y/o dependencia** (art 173.2 CP). Se entiende por violencia doméstica toda la violencia, física o psíquica, ejercida en el núcleo familiar, de manera que su ámbito se extiende a todo el círculo de personas que conviven, pretendiendo con ello otorgar una especial protección a la víctima, precisamente atendiendo a ese especial vínculo.

El registro de sesgos aquí planteado no pretende excluir los mismos del procesado, sino identificar si se tienen en cuenta a la hora de construir los algoritmos, al objeto de asegurar que se elaboran herramientas adecuadas para el tratamiento de la información.

SERVICIO	TÉCNICOS	CONTENIDO	REGISTRO
CD	<ul style="list-style-type: none"> <li>Cuantitativos: obteniendo un número adecuado de documentos a clasificar de cada tipología (unos 50 aprox.).</li> <li>Geográficos: seleccionando equilibradamente los documentos entre las tres sedes</li> <li>Temporales: selección de documentos equilibrada en el periodo temporal suministrado.</li> </ul>	<p>No se incluyen ni emplean, fuera de las determinadas por el propio mapa documental, características que tengan que ver con: edad, sexo, capacidades de las personas, cargos o profesiones, raza, nacionalidad, orientación sexual, religión, ideología, lengua, localización de residencia ni relaciones de afectividad y/o dependencia.</p> <p>Tampoco se contemplan: opiniones políticas, convicciones religiosas o filosóficas, ni datos de afiliación sindical, genéticos, biométricos o relativos a la salud.</p>	09/03/2022
NER	<ul style="list-style-type: none"> <li>Cuantitativos: obteniendo un número adecuado de etiquetas por cada clase de entidad a identificar.</li> <li>Geográficos: seleccionando equilibradamente los documentos de las tres sedes disponibles</li> <li>Temporales: selección de documentos equilibrada dentro del periodo temporal suministrado.</li> </ul>	<p>No se identifican características que tengan que ver con: edad, sexo, capacidades de las personas, cargos o profesiones, raza, nacionalidad, orientación sexual, religión, ideología, lengua, localización de residencia ni relaciones de afectividad y/o dependencia.</p> <p>Tampoco se contemplan: opiniones políticas, convicciones religiosas o filosóficas, ni datos de afiliación sindical, genéticos, biométricos o relativos a la salud.</p>	09/03/2022
SIM	<ul style="list-style-type: none"> <li>Cuantitativos: obteniendo un número adecuado de etiquetas identificadas de nombres y apellidos, así como de denominaciones sociales hasta identificar similitud entre ellas por variaciones de acrónimos, falta de algún elemento, etc.</li> <li>Geográficos: seleccionando equilibradamente los documentos de las tres sedes disponibles</li> <li>Temporales: selección de documentos equilibrada dentro del periodo temporal suministrado.</li> </ul>	<p>No se identifican características que tengan que ver con: edad, sexo, capacidades de las personas, cargos o profesiones, raza, nacionalidad, orientación sexual, religión, ideología, lengua, localización de residencia ni relaciones de afectividad y/o dependencia.</p> <p>Tampoco se contemplan: opiniones políticas, convicciones religiosas o filosóficas, ni datos de afiliación sindical, genéticos, biométricos o relativos a la salud.</p>	27/02/2023
CO	<ul style="list-style-type: none"> <li>Cuantitativos: obteniendo un número adecuado de documentos a clasificar de cada tipología (+6000 por cada Orden)</li> <li>Geográficos: seleccionando equilibradamente los documentos entre las tres sedes de las que se extrajeron los documentos.</li> <li>Temporales: selección de documentos equilibrada en el periodo temporal suministrado.</li> </ul>	<p>No se incluyen ni emplean características que tengan que ver con: edad, sexo, capacidades de las personas, cargos o profesiones, raza, nacionalidad, orientación sexual, religión, ideología, lengua, localización de residencia ni relaciones de afectividad y/o dependencia.</p> <p>Tampoco se contemplan: opiniones políticas, convicciones religiosas o filosóficas, ni datos de afiliación sindical, genéticos, biométricos o relativos a la salud.</p>	03/05/2023

SERVICIO	TÉCNICOS	CONTENIDO	REGISTRO
CR	<ul style="list-style-type: none"> <li>Cuantitativos: obteniendo un número adecuado de documentos a clasificar de cada clase Registral (+50 por cada clase)</li> <li>Geográficos: seleccionando equilibradamente los documentos entre las tres sedes de las que se extrajeron los documento.</li> <li>Temporales: selección de documentos equilibrada en el periodo temporal suministrado.</li> </ul>	<p>No se incluyen ni emplean, fuera de las determinadas por el propio mapa documental, características que tengan que ver con: edad, sexo, capacidades de las personas, cargos o profesiones, raza, nacionalidad, orientación sexual, religión, ideología, lengua, localización de residencia ni relaciones de afectividad y/o dependencia.</p> <p>Tampoco se contemplan: opiniones políticas, convicciones religiosas o filosóficas, ni datos de afiliación sindical, genéticos, biométricos o relativos a la salud.</p>	03/05/2023
RE	<ul style="list-style-type: none"> <li>Cuantitativos: obteniendo un número adecuado de entidades y de sus distintas categorías por cada relación a establecer.</li> <li>Geográficos: seleccionando equilibradamente los documentos de las tres sedes disponibles</li> <li>Temporales: selección de documentos equilibrada dentro del periodo temporal suministrado.</li> </ul>	<p>No se identifican características que tengan que ver con: edad, sexo, capacidades de las personas, cargos o profesiones, raza, nacionalidad, orientación sexual, religión, ideología, lengua, localización de residencia ni relaciones de afectividad y/o dependencia.</p> <p>Tampoco se contemplan: opiniones políticas, convicciones religiosas o filosóficas, ni datos de afiliación sindical, genéticos, biométricos o relativos a la salud.</p>	03/05/2023
SI	<ul style="list-style-type: none"> <li>Cuantitativos: anonimizado de los documentos de los cuales se va a obtener el resumen.</li> <li>Geográficos: abstrayendo del documento cualquier referencia o localización por domicilio.</li> </ul>	<p>No se identifican características que tengan que ver con: edad, sexo, capacidades de las personas, cargos o profesiones, raza, nacionalidad, orientación sexual, religión, ideología, lengua, localización de residencia ni relaciones de afectividad y/o dependencia.</p> <p>Tampoco se contemplan: opiniones políticas, convicciones religiosas o filosóficas, ni datos de afiliación sindical, genéticos, biométricos o relativos a la salud.</p>	03/05/2023

SISTEMA DE ANONIMIZADO DE DOCUMENTOS	VALORACIÓN							
Descripción	C	I	A	T	D	RTO	RPO	Categoría del sistema
Sistema de anonimizado de documentos	Alto	Alto	Alto	Alto	Medio	4 horas / 1 día	N/A	Alta

## Justificación de las valoraciones

Las valoraciones de (C), (I), (A) y (T) del activo tipo información, se basan en los niveles establecidos para el sistema Minerva (del que proviene el documento y en el que se almacenará firmado). Es información jurisdiccional de cualquier orden.

En relación al activo esencial de tipo Servicio, se valora (D) en MEDIO porque es una opción voluntaria para el ciudadano; es decir, se le ofrece también la posibilidad de firmar de manera manuscrita sobre formato papel. Por tanto, ante la indisponibilidad del sistema, habría métodos de firma alternativos.

## Confidencialidad:

-El documento judicial o fiscal electrónico objeto del proceso de seudoanonimizado tiene la misma valoración de sus diferentes dimensiones de seguridad otorgada en su sistema origen, el sistema de gestión procesal o fiscal.

-El documento resultado del proceso también constituye un documento judicial o fiscal electrónico. Sin embargo, el seudonimizado no supone el anonimizado con la absoluta desvinculación de los datos personales de sus titulares. Asimismo, pueden existir datos personales, incluso categorías especiales de datos que no hayan sido objeto del proceso porque es necesario su conocimiento por el destinatario del documento. Por ello, el nivel de confidencialidad deberá seguir siendo el mismo, ya que las consecuencias de su acceso indebido o su revelación no autorizada puede ocasionar daños irreversibles a los titulares de los datos personales, principalmente en sus derechos fundamentales al honor, intimidad personal y familiar, imagen, y protección de datos, sin perjuicio de otros derechos.

**Integridad:** El documento objeto del proceso de seudonimizado debe mantener su integridad con carácter previo al proceso.

La detección de las entidades nominales y la ejecución de las opciones de seudonimización ofrecidas al usuario, no puede afectar a la integridad del resto de los datos contenidos en el documento. La existencia de estas posibles alteraciones podría suponer un perjuicio para el propósito de su destinatario. Este perjuicio afectaría principalmente a su derecho a la tutela judicial efectiva.

## SISTEMA DE ANONIMIZADO DE DOCUMENTOS

### Autenticidad:

- a) Autenticidad de la información y de la fuente de origen: El documento objeto del proceso tiene un nivel alto de autenticidad de la información contenida en el mismo, y su origen, de conformidad con su firma electrónica incorporada, y dada su consideración como documento judicial electrónico. El documento resultado del proceso también es un documento judicial electrónico y debe mantener su nivel alto de autenticidad.
- b) Autenticidad de los usuarios del acceso al servicio y a la información: para las personas que acceden sistema de seudonimización, y por tanto, al documento original y al proceso, debe garantizarse que sean las personas autorizadas. El riesgo de suplantación de identidad del funcionario autorizado, dada la naturaleza del documento y del alto impacto en el acceso indebido a la información jurisdiccional, requiere un nivel alto de autenticidad.

Trazabilidad: En términos generales, el acceso a cualquier documento judicial electrónico debe quedar registrado con objeto de evitar que puedan cometerse infracciones por parte de los usuarios del sistema, que en caso del mero acceso a categorías especiales de datos es constitutivo de delito del artículo 197.2 del Código Penal. El conocer quién accede al servicio, a la información y a los datos personales que pudiera la misma contener, permite investigar posibles accesos y un uso indebido. Dada la naturaleza de la información contenida en los documentos judiciales y su elevado impacto en los derechos y libertades de las personas físicas obrantes en los mismos, es preciso garantizar un nivel alto de trazabilidad.

Disponibilidad: Se considera que el sistema podría estar indisponible hasta 24 horas, dada la naturaleza y uso del sistema, no resultando su indisponibilidad crítica en comparación con otros sistemas obrantes en los órganos y oficinas judiciales y fiscales.

SISTEMA DE CLASIFICADO DE DOCUMENTOS	VALORACIÓN							
Descripción	C	I	A	T	D	RTO	RPO	Categoría del sistema
Sistema de clasificación de documentos	Alto	Bajo	Alto	Alto	Bajo	1-5 días	N/A	Alta

## Justificación de las valoraciones

### Confidencialidad:

El documento jurídico objeto del proceso de clasificación dirigido a los órganos judiciales y/o fiscales puede contener categorías de especiales de datos, datos de condenas e infracciones penales, así como otros datos relativos a la esfera íntima de las personas, incluidas personas vulnerables. La pérdida de confidencialidad del documento podría afectar de manera muy grave a los derechos y libertades fundamentales de sus titulares, principalmente sus derechos fundamentales a la protección de datos personales, al honor, intimidad personal y familiar e imagen, incluso podría afectar a la tutela judicial efectiva dado su contexto.

### Integridad:

El servicio de clasificación obtiene el documento subido por el usuario, devolviendo una propuesta de clasificación con un porcentaje de éxito o de acierto. En este sentido, aunque el documento perdiera accidentalmente o de forma indebida su integridad en el proceso, no sería accesible o descargado por el usuario. El único perjuicio, en su caso, sería obtener una clasificación incorrecta, que sería subsanada bien por el propio usuario al decidir aceptar la propuesta de clasificación, o bien por el órgano u oficina judicial o fiscal de destino del documento antes de integrar el documento en el sistema de gestión procesal.

### Autenticidad:

El presente servicio es una funcionalidad integrada en otros sistemas, como LexNET, que tienen un nivel de autenticidad determinado, por lo que no se valora el nivel de autenticidad del usuario en el acceso a la funcionalidad o llamada al servicio de clasificación, que ya viene dada por el sistema principal con sus medidas correspondientes. Al ser un sistema automatizado, la autenticidad aquí valorada estaría orientada a las entidades asociadas a los eventos sobre el documento objeto de clasificación, con objeto de garantizar principalmente la confidencialidad del documento. Se establece un nivel alto de acuerdo con el nivel de confidencialidad.

## SISTEMA DE CLASIFICADO DE DOCUMENTOS

### Trazabilidad:

En términos generales, el acceso a cualquier documento judicial electrónico debe quedar registrado con objeto de evitar que puedan cometerse infracciones por parte de los usuarios del sistema, que en caso del mero acceso, sin autorización, a categorías especiales de datos es constitutivo de delito del artículo 197.2 del Código Penal. El conocer quién accede al servicio, a la información y a los datos personales que pudiera la misma contener, permite investigar posibles accesos y un uso indebido. Dada la naturaleza de la información contenida en los documentos judiciales y su elevado impacto en los derechos y libertades de las personas físicas obrantes en los mismos, es preciso garantizar un nivel alto de trazabilidad.

El presente servicio, es un proceso automatizado sin intervención de usuarios, lo que limita el riesgo al registro de las entidades intervinientes en los eventos sobre el documento. Se establece un nivel alto de acuerdo con el nivel de confidencialidad.

Disponibilidad: se propone un nivel BAJO del servicio, porque la clasificación se puede realizar de manera manual ante una indisponibilidad del mismo.

En cuanto al activo “tabla de tipos documentales y probabilidades”, se valoran todas las dimensiones a BAJO, ya que no contiene datos personales, y la información no persiste en el sistema.

SISTEMA DE CLASIFICADO DE DOCUMENTOS	VALORACIÓN							
Descripción	C	I	A	T	D	RTO	RPO	Categoría del sistema
Sistema de resumen de documentos	Alto	Alto	Alto	Alto	Bajo	1-5 días	N/A	Alta

## Justificación de las valoraciones

### Documento a resumir:

Las valoraciones de (C), (A) y (T) del activo tipo información “documento a resumir”, se basan en los niveles establecidos para documentos susceptibles de formar parte del expediente judicial o fiscal o que ya están integrados en el mismo. Es información jurisdiccional de cualquier orden. Se valora (I) en BAJO ya que el documento podría volver a subirse desde el sistema en el que esté almacenado.

### Resumen del documento:

Mismas valoraciones y justificaciones que el activo anterior, a excepción de (I) que se valora en ALTO sobre todo por las consecuencias que tendría un efecto que afecta a la integridad del resumen del tipo lectura fácil dirigido a personas con discapacidad.

### Sistema de resumen de documentos:

En relación al activo esencial de tipo Servicio, se valora (D) en BAJO porque se considera un servicio de apoyo, cuya indisponibilidad no afectaría a la tramitación de los procedimientos.

# Gracias por su atención



GOBIERNO DE ESPAÑA

MINISTERIO DE JUSTICIA

SECRETARÍA DE ESTADO DE JUSTICIA

SECRETARÍA GENERAL PARA LA INNOVACIÓN Y CALIDAD DEL SERVICIO PÚBLICO DE JUSTICIA

DIRECCIÓN GENERAL DE TRANSFORMACIÓN DIGITAL DE LA ADMINISTRACIÓN DE JUSTICIA

SUBDIRECCIÓN GENERAL DE DESARROLLO E IMPLANTACIÓN DE SERVICIOS DIGITALES



AGENDA 2030