



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

**LICENCIATURA EN ADMINISTRACIÓN
DE ARCHIVOS Y GESTIÓN
DOCUMENTAL**

ESCUELA NACIONAL DE ESTUDIOS SUPERIORES
UNIDAD MORELIA

LA FIABILIDAD DE LA INTELIGENCIA
ARTIFICIAL EN LA PROTECCIÓN DE DATOS
PERSONALES: UNA MIRADA DESDE LA
ARCHIVÍSTICA

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
**LICENCIADO EN ADMINISTRACIÓN DE
ARCHIVOS Y GESTIÓN DOCUMENTAL**

P R E S E N T A:

DANTE ISAAC VEGA SALINAS

**DIRECTOR DE TESIS:
DR. LUIS MIGUEL GARCIA VELAZQUEZ**

MORELIA, MICHOACAN

MARZO, 2024



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



ESCUELA
NACIONAL
DE ESTUDIOS
SUPERIORES
UNIDAD MORELIA

10
años
(2011-2021)

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO ESCUELA
NACIONAL DE ESTUDIOS SUPERIORES UNIDAD MORELIA
SECRETARÍA GENERAL
SERVICIOS ESCOLARES

MTRA. IVONNE RAMÍREZ WENCE

DIRECTORA

DIRECCIÓN GENERAL DE ADMINISTRACIÓN ESCOLAR

PRESENTE

Por medio de la presente me permito informar a usted que en la **sesión ordinaria 01 del Comité Académico de la Licenciatura en Administración de Archivos y Gestión Documental** de la Escuela Nacional de Estudios Superiores (ENES), Unidad Morelia, celebrada el día **17 de enero de 2024**, se acordó poner a su consideración el siguiente jurado para la presentación del Trabajo Profesional del alumno **Dante Isaac Vega Salinas** de la Licenciatura en **Administración de Archivos y Gestión Documental**, con número de cuenta **417019154**, con el trabajo titulado: **"La fiabilidad de la inteligencia artificial en la protección de datos personales: una mirada desde la archivística"**, bajo la dirección como tutor del **Dr. Luis Miguel García Velázquez**.

El jurado queda integrado de la siguiente manera:

Presidente:	Dra. Marisol Flores Garrido
Vocal:	Dra. Yaminel Bernal Astorga
Secretario:	Dr. Luis Miguel García Velázquez
Suplente:	Dra. Jennifer Sabrina Torres Martínez
Suplente:	Lic. Salvador Luna Perales

Sin otro particular, quedo de usted.

Atentamente
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Morelia, Michoacán a 03 de abril de 2024.

DRA. ANA YESICA MARTÍNEZ VILLALBA
SECRETARIA GENERAL

CAMPUS MORELIA

Antigua Carretera a Pátzcuaro N° 8701, Col. Ex Hacienda de San José de la Huerta
58190, Morelia, Michoacán, México. Tel: (443)689.3500 y (55)5623.7300, Extensión Red UNAM: 80614
www.enesmorelia.unam.mx

AGRADECIMIENTOS INSTITUCIONALES

A la Universidad Nacional Autónoma de México y la Escuela Nacional de Estudios Superiores Unidad Morelia, por brindarme la posibilidad de una educación de calidad y ser un espacio de desarrollo académico, cultural y personal.

A la Licenciatura de Administración de Archivos y Gestión Documental, por la formación y enseñarme las distintas posibilidades de la profesión archivística.

A mi asesor de tesis el Dr. Luis Miguel García Velázquez, por su confianza, tiempo, paciencia y apoyo en todo el proceso de realización de esta tesis.

A Alicia Barnard por la oportunidad de incorporarme al proyecto InterPARES Trust AI

A la Dra. Yaminel Bernal Astorga por creer que este trabajo podría ser más de lo que en un principio pretendía ser.

A la Dra. Marisol Flores Garrido por su interés y apreciación en el alcance de esta tesis.

A la Dra. Sabrina Torres Martínez por enseñarme las distintas posibilidades entre la tecnología y la archivística

Al Lic. Salvador Luna Perales, por su apoyo y apertura en mis inquietudes en esta carrera.

A mi jurado revisor, conformado por los anteriormente mencionados. Tanto sus comentarios sobre la forma como sobre el contenido fueron valiosos para la culminación de este trabajo.

AGRADECIMIENTOS PERSONALES

A mi madre Claudia Esperanza, por su incondicional amor, fe y apoyo a lo largo de toda mi vida. Aprecio y valoro profundamente el sacrificio que hiciste para otorgar esta posibilidad. Gracias por los principios y valores que me diste. Lo presentado es un logro, prueba de la esperanza puesta en mí, a través del tiempo.

A mis hermanas Naomi y Sofía, por acompañarme, motivarme, ser soporte y soportarme. Gracias por cuidar de mí, su afecto y cariño. Todo esto me alentó, durante estos largos años. Las adoro y siempre esfuércense en cumplir sus sueños.

A mi mejor amigo Bethel por siempre estar ahí, y darme esos ánimos que necesitaba. A mis estimados amigos Christian, Juan Carlos y Fernanda, gracias por los buenos momentos que pasamos juntos. Al igual que a mi amigo José.

A mis amigos de TIC's: Yovanny, Pablo Clemente, Andrés, Donovan y Josué, por brindarme un espacio en su círculo de amistad. A mi amigo Pablo Rafael, por su escucha y ayuda.

A la abuelita Rita y las tías Claudia, Porfirio y Bárbara, por su amabilidad y atención.

Agradezco a todas aquellas personas, compañeros y maestros que, con buena voluntad, me brindaron su apoyo.

En última instancia, deseo dedicar este trabajo a quienes sufren de violencia y atraviesan adversidades, que este trabajo sea testimonio de la probabilidad de éxito en la incertidumbre de las circunstancias. La dignidad no es una moneda de cambio, y siempre debe lucharse por conservarla. De lo contrario ¿Qué nos quedaría?

Resumen

Esta tesis explora el estudio teórico de la fiabilidad de la Inteligencia Artificial (IA) en el contexto de la protección de datos personales y el papel de los archivistas en su uso y responsabilidad. Se trata de una investigación descriptiva basada en una revisión de literatura, cuyo objetivo es respaldar un estudio de caso dentro del proyecto multinacional InterPARES TRUST AI. La investigación pretende identificar antecedentes desde la IA y la archivística que sirvan de punto de partida para la identificación y extracción automatizada de datos personales y sensibles. Se presentan conceptos fundamentales de archivo y documento, conceptos de IA y sesgos algorítmicos. Después se recorren casos de estudio que demuestran el uso de la IA en el ámbito archivístico, dando como resultado un marco de acción en la interacción entre ambas disciplinas. Para finalizar, se aborda el contexto de aplicación de la IA en la protección de datos personales para el caso particular de la anonimización, delineando también acciones concretas para el perfil profesional archivístico en el uso de la IA.

Palabras clave: Archivística, Inteligencia Artificial, Datos personales, Revisión de literatura, Anonimización

Abstract

This thesis explores the theoretical study of the reliability of Artificial Intelligence (AI) in the context of personal data protection and the role of archivists in its use and responsibility. It is a descriptive research based on a literature review, which aims to support a case study within the multinational InterPARES TRUST AI project. The research aims to identify antecedents from AI and archival science that serve as a starting point for the automated identification and extraction of personal and sensitive data. Fundamental archival and document concepts, AI concepts and algorithmic biases are presented. Case studies demonstrating the use of AI in the archival domain are then reviewed, resulting in a framework for action in the interaction between the two disciplines. Finally, the context of application of AI in the protection of personal data for the particular case of anonymization is addressed, also outlining concrete actions for the archival professional profile in the use of AI.

Keywords: Archival science, Artificial Intelligence, Personal data, Literature review, Anonymization

Contenido

Introducción	7
Búsqueda de literatura	18
Planteamiento y estrategia de búsqueda	18
Organización y selección de los documentos	23
Redacción de la revisión de literatura	24
Reflexiones del proceso de búsqueda de literatura	25
El documento de archivo y la inteligencia artificial	26
El archivo y el documento de archivo.....	26
El documento de archivo digital y electrónico.....	30
Los conceptos de Inteligencia Artificial y Aprendizaje Automático	32
Sesgos Algorítmicos.....	36
El panorama profesional del archivista en el uso de la inteligencia artificial en el archivo	40
El abordaje de la IA en el archivo	40
Explicabilidad e interpretabilidad de la IA en tratamiento de la información	40
El contacto del archivista con la IA	46
Una gestión archivística de los datos.....	48
Principios de diseño para un software de procesamiento de documentos	51
Casos de estudio en el uso de la IA en los archivos	53
La protección de datos personales frente a la inteligencia artificial y su fiabilidad desde la archivística	70
Las normativas de transparencia y privacidad en México	70
Los datos personales, las versiones públicas y la anonimización	74
La inteligencia artificial y la anonimización	85
La fiabilidad de la inteligencia artificial en la anonimización de datos personales.....	91
Consideraciones finales.....	98
Referencias.....	104
ANEXO.....	110

Índice de Tablas

Tabla 1. Dimensiones de estudio de InterPARES 2	10
Tabla 2. Estrategias de búsqueda Archivística e Inteligencia Artificial	20
Tabla 3. Revistas académicas consultadas.....	21
Tabla 4. Estrategias de búsqueda Eje Regulación y tecnología.....	22
Tabla 5. Ejemplo de organización de la literatura seleccionada.....	24
Tabla 6 Propuesta de Flujo de Valoración para datos.	50
Tabla 7. Funciones y acciones del archivista en relación con la IA en general.....	67
Tabla 8. Funciones y acciones del archivista en relación a la fiabilidad de la IA en la protección de datos personales.....	97

Introducción

En la actualidad del siglo XXI las tecnologías de la información se integran a las organizaciones en una mayor cantidad de ámbitos. Esto ha ampliado el modo de producción y resguardo de los documentos en formato digital¹, transformando a su vez al archivo al insertarse en un ambiente de alta conectividad; en consecuencia, surgen nuevas necesidades de estudio acerca de las implicaciones en torno a la archivística y al archivo.

La Inteligencia Artificial (IA) es un área de desarrollo científico y tecnológico que en los últimos años ha tenido un mayor uso al extender sus posibilidades funcionales. El archivo es un espacio que poco a poco está formando parte de su ámbito de implementación, dado que el uso de la IA, ofrece la capacidad tecnológica de procesar grandes cantidades de datos o información, para su gestión y estudio.

En el contexto de una disrupción tecnológica como lo es la utilización de la IA y los archivos, el proyecto *International Research on Permanent Authentic Records in Electronic Systems InterPARES*, en su fase *Trust AI*, coordina actualmente una serie de investigaciones en torno al desarrollo responsable de la IA para realizar diseños y desarrollos que aprovechen la inteligencia artificial como medio de soporte y ayuda para la disponibilidad y acceso a los documentos. Para lo anterior ha generado asociaciones con la academia, los archivos, el gobierno y la industria. El marco de referencia utilizado para guiar estas investigaciones lo proveen la archivística y la diplomática ("*InterPARES Trust AI - Artificial Intelligence*" 2021).

Es necesario profundizar en el proyecto *InterPARES* dada la relevancia de su trayectoria, que ha generado un desarrollo teórico particular de especial utilidad para entender la relación entre tecnología y archivística. El antecedente directo más cercano fue el *Preservation of the Integrity of Electronics Records* (UBC Project), desarrollado entre los

¹ El presente trabajo está alineado con la terminología del Glosario de preservación archivística digital versión 4.0 de Juan Voutssas-M. y Alicia Barnard Amozorrutia. Para este término u otros similares consultarlo en la siguiente dirección: https://iibi.unam.mx/archivistica/glosario_preservacion_archivistica_digital_v4.0.pdf

periodos de 1994-1997, dirigido por Luciana Duranti y Terry Eatwood; en tal proyecto se buscaba estudiar la integridad de los documentos electrónicos, pero a partir de la perspectiva de los creadores de los documentos (Mena Mugica, 2020). La metodología del proyecto fue deductiva, con base en los conceptos de la diplomática (Mena Mugica, 2020). Más adelante, se creó el proyecto *InterPARES I*, durante el periodo de 1999-2001, con el objetivo de estudiar la preservación a largo plazo y la autenticidad del documento electrónico de archivo.

La investigación en esta fase se centró en varios dominios² de estudio, que fueron elaborados de la siguiente forma:

- Autenticidad: Identificar los requerimientos para la evaluación y mantenimiento de la autenticidad de los documentos electrónicos.
- Valoración: Establecer si la metodología de valoración de documentos que tradicionalmente se utilizaba era aplicable, considerando las tecnologías digitales.
- Preservación: Establecer métodos para la preservación de documentos electrónicos auténticos.
- Estrategia: La generación de la normativa a nivel internacional, nacional y organizacional para la preservación a largo plazo de documentos electrónicos auténticos

La finalidad de esta iniciativa fue proponer elementos para la creación de normativas en los niveles internacional, nacional y organizacional, que atendieran la preservación a largo de documentos electrónicos auténticos. La investigación realizada por InterPARES se cristalizó en varios productos³ entre los cuales se hallan:

- El documento *Plantilla de Análisis*, que identifica todos los elementos conocidos de un expediente electrónico basándose en los principios y conceptos de la diplomática y la archivística.

² Para mayor información consultar: http://www.interpares.org/ip1/ip1_goals&methodologies.cfm

³ Para mayor información consultar: http://www.interpares.org/ip1/ip1_documents.cfm?cat=atf

- Listados de requisitos para evaluar y mantener la autenticidad de los documentos de archivo electrónicos.
- Documentos que comparan los requisitos de autenticidad de *InterPARES* con las normas existentes en ese momento.

En el contexto de este proyecto surgió la necesidad de investigar requisitos para aquellos documentos que alcanzan el estado de permanente, a diferencia del UBC Project, enfocado más en los documentos activos y semiactivos.

A continuación, se abordará de manera breve el objetivo de las siguientes fases del proyecto, para así enmarcar la fase etapa actual Trust AI - relacionada directamente con este trabajo- como continuación de las anteriores fases.

La siguiente fase del proyecto, titulada como *InterPARES 2 Project: Experiential, Interactive, Dynamic Records*, se desarrolló entre 2002 y 2007; el objetivo de esa etapa fue el estudio de los documentos experienciales, interactivos y dinámicos, con el fin de crear la serie de principios necesarios para asegurar la fiabilidad y exactitud, además de la preservación a largo plazo de los documentos generados por medios computacionales. Este enfoque derivó de la preocupación de la estabilidad de los documentos digitales, dada su fragilidad, obsolescencia, manipulación y posible incompatibilidad, para así asegurar el acceso a la memoria a futuras generaciones (InterPARES 2 Project, 2007).

En esta fase los dominios de estudio fueron los que se indican en la Tabla 1.

Dominio		Ideas principales
Dominio de Creación de documentos mantenimiento.	1. de y	<ul style="list-style-type: none"> • La caracterización de documentos de archivo que necesitan estar en línea para su consulta. • La posible alteración del proceso de creación de los documentos. • La identificación de los elementos formales de los documentos tanto en formato físico como digital.
Dominio de fiabilidad, autenticidad y precisión de los documentos.	2. La y	<ul style="list-style-type: none"> • El entendimiento del significado de fiabilidad en contextos artísticos, científicos y gubernamentales. • La delimitación del alcance del concepto de precisión durante la creación del documento en cada caso de estudio. • Reconocer la autenticidad en el contexto de cada documento.
Dominio de métodos de valoración y preservación.	3. Los de y	<ul style="list-style-type: none"> • La aplicabilidad de los resultados de InterPARES 1 a los contextos estudiados en InterPARES 2.

Tabla 1. Dimensiones de estudio de InterPARES 2 Fuente: Elaboración propia

Por otro parte, en esta fase de *InterPARES 2* se consideraron los dominios de cruce (*cross domain*), donde se implicó la interacción de varias áreas de estudio. Estos fueron:

- Descripción⁴: Se estudió la creación, captura, gestión y preservación de metadatos.
- Políticas⁵: Se evaluaron aquellas normas y procedimientos para la creación, mantenimiento, preservación y uso de documentos, además de la formulación de criterios para la generación de nuevas políticas.
- Terminología⁶: Se estableció si los términos utilizados en el proyecto eran adecuados y se delimitó su relación con otras áreas.

El proyecto *InterPARES 3*, titulado como *Theoretical Elaborations into Archival Management (TEAM): Implementing the theory of preservation of authentic records in digital systems in small and medium-sized archival organizations*, se realizó entre 2007 y 2012; su objetivo fue elaborar planes de acción para aplicar medidas de preservación en documentos digitales para el caso de instituciones pequeñas y medianas (InterPARES 3 Project, 2012).

⁴ Para mayor información consultar: http://www.interpares.org/ip2/ip2_description.cfm

⁵ Para mayor información consultar: http://www.interpares.org/ip2/ip2_policy.cfm

⁶ Para mayor información consultar: http://www.interpares.org/ip2/ip2_terminology.cfm

El proyecto *InterPARES Trust*⁷, llevado a cabo desde 2012 hasta 2018, tenía como objetivo el estudio de la confianza y confiabilidad de los datos y documentos en ambientes línea, para la creación de metodologías que pudieran asegurar su buena gobernanza, obtener la confianza del público y producir una memoria digital persistente (InterPARES Trust, 2018).

Los dominios de investigación fueron los siguientes:

- Infraestructura: Aquella infraestructura que da alojamiento a los documentos en entornos digitales en línea.
- Seguridad: Los problemas de seguridad en relación con los datos alojados en línea.
- Control: La gestión de los elementos de fiabilidad, exactitud y autenticidad de los documentos en línea.
- Acceso: Los datos abiertos y el derecho a ser olvidado, entre otras temáticas.
- Legal: La cadena de evidencia y su autenticación, entre otros aspectos.

Los dominios de investigación cruzados fueron los siguientes;

- Terminología: La producción de un glosario multi-lenguaje que explique el uso de los términos dentro del proyecto.
- Recursos: La constante producción de bibliografía anotada que identifique artículos relevantes para la investigación.
- Educación: La identificación de aquellos modelos que sirvan para la transmisión del conocimiento generado a partir de la investigación.
- Política: La creación de políticas de gestión para cada uno de los dominios de la investigación.
- Social/Problemas Sociales: El entendimiento de cómo la información que se halla en internet puede generar problemas sociales o influir en la sociedad.

⁷ Para mayor información consultar: http://interparestrust.org/trust/about_research/domains

A lo largo de más de 20 años, el proyecto *InterPARES* se ha desplazado desde la aplicación de la diplomática en el documento electrónico, considerando los elementos centrales de fiabilidad, autenticidad y exactitud, hacia el estudio de su gestión en ambientes digitales y dentro de un entorno en línea. Cabe destacar la larga trayectoria y el alcance que *InterPARES* ha realizado a través del tiempo. Como consecuencia de la mayor cercanía entre la inteligencia artificial y su aplicación en los archivos, surge la siguiente fase del proyecto *InterPARES*, titulada como *Trust AI*, que está planeada para suceder en el periodo comprendido entre 2021 y 2026. El objetivo principal para esta etapa es establecer los conceptos necesarios para los desarrollos responsables de la IA en los archivos con base en principios archivísticos y de la diplomática.

Los proyectos bajo el amparo de *InterPARES* regularmente llevan a cabo una revisión de literatura previa, que ofrezca contexto para cada investigación a realizar; la que se incluye en el presente trabajo se propone situar la discusión de Inteligencia Artificial y Archivística para orientar el proyecto *Case Study on Extraction and Identification of Records containing Personal Data and Sensitive Personal Data for Long Term Preservation (RA02, InterPARES Trust AI)*, realizada con la asesoría del Dr. Luis Miguel García Velázquez, quien ha sido colaborador en este grupo de investigación.

En la descripción de dicho proyecto se declara que:

El objetivo de este estudio es desarrollar un algoritmo para reconocer y extraer información no estructurada que sean datos personales y datos personales sensibles en documentos digitalizados (PDF con OCR) mediante la aplicación de técnicas de inteligencia artificial (IA), en particular algoritmos de aprendizaje automático (AM), y buscar posibles requisitos o equivalentes de confiabilidad (exactitud, fiabilidad y autenticidad) de la IA y el AM del producto que se va a obtener. (*InterPARES Trust AI*, 2021).

Cabe señalar que el proyecto fue realizado por un equipo de investigación interinstitucional, con participación de la Universidad Nacional Autónoma de México, el Colegio de México, el Instituto Nacional de Transparencia, Acceso a la Información y

Protección de Datos Personales (INAI) y la consultoría especializada, quienes acordaron que el caso de estudio debería desarrollarse en el contexto de las instituciones públicas.

Así, el problema de investigación en el que se inserta este trabajo responde a la necesidad de proveer un punto de partida que integre las perspectivas de la archivística y la IA para proyectar un desarrollo semejante, a sabiendas que la complejidad de protección de datos personales responde a una creciente demanda de información por parte de las comunidades usuarias, así como de los requerimientos en términos de transparencia que se desprenden de las normativas vigentes. La extracción automatizada de datos personales debe considerar las prácticas archivísticas ya establecidas para el testado de documentos, al tiempo que proponen mecanismos de automatización que se enfrenten a los crecientes volúmenes documentales que gestionan las distintas instituciones.

Mediante una investigación en artículos científicos nos hemos propuesto abordar la siguiente pregunta de investigación: ¿Cuáles son aquellos antecedentes desde la IA y la archivística que ofrecen un punto de partida, recuperando aprendizajes de experiencias previas, para servir de sustento conceptual, metodológico y empírico que sea relevante para la identificación y extracción automatizada de datos personales e información sensible?

Para dar cauce esta investigación descriptiva, hemos establecido los siguientes objetivos:

Objetivo General

Realizar una revisión de literatura con apego al objetivo del caso de estudio *Case Study on Extraction and Identification of Records containing Personal Data and Sensitive Personal Data for Long Term Preservation (RA02)*, desde una perspectiva archivística, que resulte pertinente para el entendimiento de las posibilidades y las consecuencias de los desarrollos responsables de inteligencia artificial en el área, así como del potencial de la profesión de la archivística para participar activamente en ellos.

Objetivos específicos

- Realizar una búsqueda de literatura que relacione la archivística, los archivos y la IA en apego más cercano al objetivo del caso de estudioal desarrollo de algoritmos para el tratamiento de datos personales, con énfasis en la confiabilidad.
- Establecer criterios de elegibilidad para la organización de la literatura recuperada con intención de integrar una visión panorámica sobre los desarrollos responsables de inteligencia artificial en los archivos, y potenciar el involucramiento activo de la profesión archivística.
- Redactar un informe que sintetice de forma breve la literatura seleccionada, de modo que sea posible establecer un marco de interacción entre el archivista y la IA

Para acercarse a la revisión y análisis de la literatura se han tomado como referentes teóricos los propuestos en el trabajo de Alicia Barnard y Juan Voutsass (2014) quienes, además de participar en las distintas iteraciones del proyecto InterPARES, han realizado esfuerzos notables por enmarcar la discusión de los archivos digitales y el aprovechamiento tecnológico en México y América Latina; particularmente, Alicia Barnard ha promovido durante los últimos años buena parte de la discusión sobre la IA y sus posibilidades de aplicación en los archivos de nuestro país, tomando como punto de partida este trabajo previo.

La obra de Barnard y Voutssas (2014) ha sido de gran provecho para vincular los distintos artículos revisados, resultando especialmente útil la conceptualización de las nociones archivísticas y tecnológicas realizada por estos autores -a partir de lo propuesto en *InterPARES* al contextualizarla para el escenario hispanoparlante, sin dejar de ocuparse para que dichas elaboraciones teóricas sean pertinentes para su consideración en la situación contemporánea de la archivística en México.

En el caso de las concepciones teóricas de la IA, se han tomado como básicas las nociones establecidas por Peter Norvig y Stuart J. Russell (2009), aunque estas han sido

complementadas a lo largo del presente trabajo con caracterizaciones de otros autores; lo anterior con la intención de proveer mayor claridad y cercanía con el contexto archivístico.

Si bien la noción de desarrollos responsables es parte de las caracterizaciones distintivas del proyecto InterPARES, para delimitar sus alcances teóricos hemos seguido dos categorizaciones propuestas sobre sesgos algorítmicos, tanto las establecidas por Batya Friedman y Helen Nissenbaum (1996) como las conceptualizadas por David Danks y Alex John London (2017); estos trabajos se recuperan no solamente porque establecen vínculos claros con las etapas de desarrollo de un modelo computacional, sino porque permiten articular dichas fases con acciones propias de la profesión archivística, lo que nos habilita para seguir las pautas conceptuales propuestas por Timnit Gebru (2020), que resultan pertinentes para nuestros propósitos.

La hipótesis a perseguir durante esta investigación es que es posible recuperar claves desde los antecedentes teóricos y los casos de estudio identificados en la literatura, que permitan establecer líneas para enlazar la orientación futura de la profesión archivística con la generación de una IA responsable.

En cualquier investigación se requiere de una metodología que brinde la rigurosidad al proceso. La revisión de literatura se puede relacionar con la metodología cualitativa descriptiva; al respecto, Sampieri (2014) comenta que la revisión de literatura tiene la utilidad en la detección de ideas y conceptos clave, además de identificar otras metodologías, lo que permite generar categorías y profundizar en la interpretación.

El contexto de la investigación se acerca a la definición de revisión literatura de alcance -en inglés, *scoping review*-, la cual consiste, según Peters et al. (2015), en un medio de reconocimiento para establecer definiciones y límites conceptuales, al igual que temas poco explorados.

A partir de lo anterior es posible identificar que la relación entre archivos y algoritmos de inteligencia artificial ha sido poco explorada en comparación con otras tecnologías, resulta pertinente partir de un enfoque de reconocimiento de la literatura disponible que pueda servir de marco referencial para el caso de estudio.

Para estructurar el proceso de revisión de literatura se tomarán como punto de partida las 5 fases propuestas por Arnau Sabatés y Sala Roca (2020).

A continuación, se presenta una descripción metodológica a partir de esta propuesta:

1. Diseñar la estrategia de búsqueda: Se localizará aquella información que servirá como base para realizar la revisión de literatura. En particular se sugiere hacer uso de palabras clave y operadores booleanos (AND, OR, NOT) en buscadores para bases de datos especializadas, definiendo en este paso los criterios de selección.
2. Seleccionar y presentar los resultados de búsqueda: Para el caso donde la revisión de literatura es el fin mismo de la investigación, se deben registrar las estrategias de búsquedas y los artículos elegidos. Con esta intención, se deben generar criterios de inclusión y exclusión, pudiendo ser estos temáticos, geográficos, metodológicos o poblacionales, entre otros. También es recomendable registrar el proceso de refinamiento en el proceso de selección de los resultados
3. Almacenar y registrar los resultados de búsqueda: En primera instancia se recomienda usar un gestor de referencias para los artículos seleccionados. Para el registro de los resultados de la búsqueda es conveniente construir una tabla con los resultados obtenidos y generar campos de información relevante.
4. Organizar los resultados de búsqueda: En este apartado se clasifican los resultados según temática, cronología, enfoque teórico, entre otros. Uno de los métodos recomendados son el de mapeo o el de índices.
5. Analizar, interpretar y discutir los resultados: A partir de los resultados de búsqueda obtenidos, se estructura el análisis de las referencias, aspirando a generar comparaciones, contrastes y divergencias, así como a tratar de establecer las relaciones posibles con el tema de estudio.

El proceso anteriormente expuesto servirá como eje principal para la realización de la revisión de literatura.

El presente documento está organizado en cuatro partes:

- En la primera sección se presenta el proceso de revisión de la literatura, junto con un breve marco de referencia para situar la discusión sobre IA en los archivos.
- En la segunda sección conceptos clave de archivística e inteligencia artificial junto con sesgos algorítmicos
- En la tercera sección aplicaciones de la IA en los archivos y su impacto en el escenario profesional
- En la cuarta sección la descripción del contexto de automatización con IA para la protección de datos personales.
- Al final de la tercera y cuarta sección se sugieren líneas de acción para la profesión archivística, agrupadas a funciones, que a partir de la revisión de literatura se perfilan como respuestas a la transformación del entorno profesional.

Búsqueda de literatura

Planteamiento y estrategia de búsqueda

La revisión de literatura se realizó durante los meses de marzo y abril de 2022. El rango abarcó desde el año 2010 hasta el mes de abril de 2022. La temática estuvo centrada en archivística, gestión documental o acercamientos desde áreas próximas, como los estudios de la información; esto porque es regular que en las revistas académicas relacionadas a la archivística se publiquen investigaciones de áreas relacionadas. Se tomaron en cuenta artículos que tuvieran un acercamiento principalmente conceptual acerca del aspecto regulatorio de la IA. El rango temporal se delimitó con la intención de identificar más artículos relativos a tecnología y archivística que pudieran coadyuvar cómo marco teórico o referencial.

Criterios de exclusión

No se incluyeron artículos que:

1. No tratan acerca de tecnología, archivística, gestión documental o legislación, ni de manera práctica ni como acercamiento teórico.
2. Fueron técnicos en su totalidad.

Estrategia de búsqueda

Se utilizó el buscador⁸ de bibliotecas UNAM (Universidad Nacional Autónoma de México). Se tuvieron dos ejes principales basados en el objetivo de la investigación, que se derivan de la naturaleza interdisciplinaria del contenido de los artículos regularmente hallados. Dichos ejes son:

⁸ Disponible en: <https://dgb.unam.mx/>

1. **Archivos, archivística, gestión documental e IA:** Este apartado incluye aquellas referencias que involucren los archivos y la IA. Dado que es una relación que ha sido poco explorada, también se tomaron aquellas referencias que, dentro del área de los archivos y la archivística, puedan ayudar a comprender la complejidad en la que se enmarca dicha relación.
2. **Regulación y tecnología:** Este apartado tiene como punto de partida las nociones confianza, fiabilidad y, en la medida de lo posible, de exactitud, junto con las posibles regulaciones aplicables hacia la protección de datos. Estos conceptos responden de manera particular a cada disciplina. Este rubro a su vez tiene acercamientos referentes a protección de datos personales y la IA, para así tratar de conceptualizar una confianza específica entre la IA y la protección de datos personales y sensibles.

Para ambos ejes se utilizó el símbolo asterisco *, que le permite al buscador EBSCO recibir un elemento raíz de una palabra como: Archiv. Esto brinda la posibilidad de de forma automática mostrar los resultados que incluyan las derivaciones como archivos, archivística, archivología y archivo. Además de “*” otros comandos empleados fueron booleanos, como AND NOT y OR.

Archivos, archivística, gestión documental e IA

Para el campo título se buscaron los términos *Archiv** e *Inteligencia Artificial*. Para el desarrollo de las subsecuentes búsquedas se siguió la lógica de combinar términos vinculados con la IA como *Aprendizaje Automático* -o en inglés *Machine Learning*- y conceptos como *Archivística* o *Archival Science* en inglés, así como *Archivo* o *Archive* en inglés. De esto, se realizaron combinaciones en los campos de búsqueda título y abstract. En la Tabla 2 que se muestra a continuación, se comparten más estrategias de búsqueda

Campos	Termino	Búsqueda
Título	Archiv*	Archiv* AND Inteligencia Artificial
Título	Archiv*	Archiv* AND IA
	IA	
Título	Archiv*	Archiv* AND Artificial Intelligence
	Artificial Intelligence	
Título	Archiv*	Archiv* AND AI
	AI	
Título	Archiv*	Archiv* AND Machine learning
	Machine learning	
Resumen	Archiv	Archiv* AND Artificial Intelligence
Resumen	Artificial Intelligence	
Término Temático	Archival Science	
Resumen	Archiv*	Archiv* AND Machine Learning
Resumen	Machine Learning	
Término temático	Archival Science	
Resumen	Archiv*	Archiv* AND Machine Learning
Resumen	Machine Learning	
Término temático	Archival Science	

Tabla 2. Estrategias de búsqueda Archivística e Inteligencia Artificial Fuente: Elaboración propia

A partir de la escasez de artículos relevantes se optó por una búsqueda manual más exhaustiva. Dada la heterogeneidad de los términos utilizados para nombrarlos, se decidió revisar los índices de las revistas de archivística en inglés a nivel internacional, puesto que en español fueron aún más escasos los resultados de búsqueda.

El rango se estableció a partir del periodo del 2010 en adelante, conforme la suscripción de la UNAM a cada revista lo permitiera. Así, en la revisión de los índices pudieron hallarse estudios relacionados en el periodo comprendido, a la par de incorporar algún término relevante que apareciera en los títulos o *abstract* que en la revisión manual se fueron hallando. En la Tabla 3 se muestran los nombres de las revistas académicas revisadas junto con el rango de búsqueda asociado para cada una de ellas.

Revista	Periodo
American Archivist	2010-2018
Journal of Archival Organization	2010-2022
Archival Science: International Journal on Recorded Information	2010-2022
Archivaria	2010-2022
Archives and Manuscripts	2012-2022
Archives and Records	2013-2022
COMMA (International Council on Archives)	2015-2021
Journal of Records Management	2010-2022

Tabla 3. Revistas académicas consultadas. Fuente: Elaboración propia

Mediante la revisión de índices se preseleccionaron 45 artículos a partir del título y el abstract. Después de una lectura completa se seleccionaron 14 artículos. Para el caso de regulación y tecnología se optó por tomar los conceptos clave de *reliability* y *trustworthiness* o en español *fiabilidad* y *confianza*. Estos a su vez fueron combinados con términos de IA, de modo que se pudiera identificar si desde otras áreas se habría estudiado la confiabilidad con un enfoque algorítmico.

En la Tabla 4, se ejemplifican las estrategias de búsqueda con la intención de aportar una mayor especificidad sobre ellas:

	Término	Búsqueda
Título	Data anonymization	Data anonymization AND Machinelearning
Título	Machine learning	
Título	Data anonymization	Data anonymization AND Machinelearning
Resumen	Machine learning	
Título	Data privacy	Data privacy AND machine learning
Resumen	Machine learning	
Abstract	Data privacy	Data privacy AND Machine learning
Abstract	Machine learning	
Titulo	Data protection	Data protection AND machine learning
Titulo	Machine learning	
Titulo	Data protection	Data protection AND machine learning
Titulo	Machine learning	
Abstract	Personal data	Personal data AND machine learning
Abstract	Machine learning	
Titulo	Trust*	Trust* AND artificial intelligence
Titulo	Artificial intelligence	

Tabla 4. Estrategias de búsqueda Eje Regulación y tecnología Fuente: Elaboración propia

Organización y selección de los documentos

A partir de los resultados de la búsqueda descrita anteriormente, se realizó una lectura rápida de los documentos, que implicaba la lectura del *abstract* y el cuerpo del documento de una manera superficial. Así fue posible identificar los distintos niveles de cercanía. Posteriormente, con una lectura profunda, se determinó si eran descartados o ubicados en cada nivel. Partiendo de la temática innovadora de la investigación y de que el contexto del caso de estudio es una entidad gubernamental, se establecieron 3 niveles de cercanía al proyecto de investigación:

- **Considerable:** El artículo trata dos o más conceptos de la investigación (en el campo de título o en el *abstract*) y el contexto (institución pública) es cercano.
- **Cercana:** El artículo trata al menos un concepto relacionado, pero el contexto puede estar alejado en relación con el del caso de estudio.
- **Aproximada:** El artículo no trata directamente con ningún concepto, pero puede servir como base teórica que proporcione contexto para la investigación o puede tener alguna implicación tecnológica relacionada.

En la Tabla 5⁹ se muestra el modo de organizar los artículos seleccionados en un primer momento, la cual utilizó los criterios nivel de cercanía anteriormente descrita como filas; además, organizó por columnas los dos ejes.

⁹ La selección final de los artículos se halla en el Anexo 1.

Ejes Temáticos

<i>Relevancia</i>	Archivos, archivística, gestión documental e IA	Diplomática, regulación, y tecnología
<i>Considerable</i>		
<i>Cercana</i>		
<i>Aproximada</i>		

Tabla 5. Ejemplo de organización de la literatura seleccionada. Fuente: Elaboración propia

Redacción de la revisión de literatura

Desde el principio fue claro que el volumen de resultados para la revisión no sería muy grande, además de que el objetivo de investigación implicaría un enfoque multidisciplinario. Se realizó un primer acercamiento mediante la realización de fichas bibliográficas, para así establecer una visión general. Este abordaje, aunque sencillo, implicaba o estaba orientado a un análisis de textos más reflexivo que permitiera, a pesar de la reducida selección de artículos, generar el mayor aprovechamiento de estos.

Al término de la generación de las fichas, ya con una visión panorámica de la literatura reunida, se procedió a iniciar la redacción del informe. Los ejes temáticos, fueron la base de dos capítulos de análisis en este informe. El primer eje motivó el capítulo “El panorama profesional del archivista en el uso de la inteligencia artificial en el archivo”, que trata de las experiencias previas donde el papel del archivista es tomado en cuenta, junto con el rol que éste puede jugar para la prevención de sesgos algorítmicos. El segundo eje dio pauta para el capítulo “La protección de datos personales frente a la inteligencia artificial y su fiabilidad desde la archivística”, que se enfoca en la anonimización automatizada a través de un breve acercamiento a la protección de datos personales.

Durante este proceso se identificaron actividades en las que el archivista pudiera verse involucrado para el desarrollo, implementación y uso de la IA -en particular, el aprendizaje automático. Estas actividades se recopilaron en una tabla, y posteriormente se agruparon en funciones. A partir de esta se identificaron funciones que el archivista puede cumplir en relación con la IA. Después se realizó una segunda tabla en la que, a partir del marco generado anteriormente, con base en las funciones se establecen una serie de acciones concretas para que el archivista pueda investigar la fiabilidad de la IA en la protección de datos personales, particularmente para la anonimización de datos personales.

[Reflexiones del proceso de búsqueda de literatura](#)

Fue posible reconocer varias fuentes documentales de gran valor acerca del estudio de la archivística. Así mismo la recopilación y selección de artículos, permitió identificar distintas metodologías que podrían ser aplicables para realizar investigación en archivos. Adicionalmente, abordar la temática de la IA permitió establecer un conocimiento general que sirve de base para su estudio de un mejor modo. Las líneas de acción generadas permiten crear una base para que futuras investigaciones puedan enriquecerla, o en su caso criticarla, para la construcción de un marco conceptual sólido que involucre la IA y los archivos.

El abordaje de una revisión de literatura puede parecer en primera instancia algo monótono o en algunos casos no tan importante. Pero es todo lo contrario, más allá de ser un proceso lineal, es uno sumamente recursivo, es decir, en investigaciones relacionadas con temas poco tratados anteriormente, se tienen que elaborar varias estrategias de búsqueda, tener un análisis crítico acerca de cada una de las fuentes de información y, más aún, reflexionar profundamente en la generación de nuevo pensamiento. Esta clase de trabajo, más que uno sencillamente de búsqueda y recuperación de información necesitó un cuidadoso análisis de textos.

El documento de archivo y la inteligencia artificial

El archivo y el documento de archivo

La Archivística es una ciencia que puede definirse de varios modos. Heredia la describe como “la ciencia que estudia la naturaleza de los archivos, los principios de su conservación y organización y los medios para su utilización” (Heredia, 1991, p.30). También puede entenderse según Mundet como “una ciencia emergente cuyo objeto son los documentos y los archivos, cuya finalidad es almacenar información y hacerla recuperable para su uso, y cuyo método se articula en torno a las normas archivísticas.” (Mundet, 2011, p.17). Ambos términos ofrecen un panorama de la Archivística que incluye elementos de conservación, organización y recuperación de la información. Con el paso del tiempo, ha cambiado la forma de considerar la naturaleza de los documentos, para dar paso a la incorporación de lo digital: en respuesta a ello se están ampliando las legislaciones, así como también se generan capacidades tecnológicas en el archivo para brindar un mejor servicio al usuario.

La Archivística se halla en una constante investigación interdisciplinar. Una aproximación de Archivística más pertinente en este contexto, nos la brinda Fuster Ruiz, al enunciar lo siguiente:

Archivística es la ciencia que se ocupa de los archivos en sus aspectos teóricos y prácticos, estableciendo principios inalterables y estudiando técnicas adecuadas de gestión de documentos, administración y tratamiento técnico de archivos, así como la función jurídica, administrativa y científica de los mismos, desde un punto de vista archivístico o de ciencias y técnicas diversas, y su relación con las entidades productoras de los conjuntos orgánicos de documentos, a fin de manejar y hacer accesible la información de los fondos documentales. (Fuster Ruiz, 1999, p. 117)

La Archivística cumple múltiples funciones sobre los archivos, como la conservación, la organización y la difusión, por mencionar algunas líneas de acción alineadas a las conceptualizaciones anteriormente citadas. De manera formal el archivo admite varias conceptualizaciones. Retomando a Heredia, ella nos dice:

Archivo es uno o más conjuntos de documentos, sea cual sea su fecha, su forma y soporte material, acumulados en un proceso natural por una persona o institución pública o privada en el transcurso de su gestión, conservados, respetando aquel orden, para servir como testimonio e información para la persona o institución que los produce, para los ciudadanos o para servir de fuentes de historia (Heredia, 1991, p.89)

Por su parte, Mundet entiende al archivo como

Un sistema corporativo de gestión que contribuye de manera efectiva, mediante una metodología propia, a la definición de los procesos de producción administrativa, garantizando la correcta creación de los documentos, su tratamiento, conservación, acceso y comunicación. (Mundet, 2011, p. 28)

La caracterización propuesta por Heredia está más referida a la procedencia de los documentos y su utilidad informativa. En el caso de Mundet, se prioriza el aspecto administrativo desde su producción hasta su difusión, una visión que en la actualidad se halla cercana a entender al archivo desde la normatividad y la gestión orientada a la consulta.

Cabe hacer una distinción entre los conceptos documento y documento de archivo. Para el caso del primero, Heredia nos indica que “documento es el testimonio de la actividad del hombre fijado en un soporte perdurable que contiene información” (Heredia, 1991, p.122). Esta definición reconoce que hay una amplia variedad de soportes en las que el ser humano deja registro de su actividad, como impresos, fotografías, mapas, discos de vinilo, casetes, tablillas o pinturas, por mencionar algunas.

Para el segundo caso, que se refiere a los documentos de archivo, cabe destacar una característica fundamental que Heredia enfatiza al describirlos como “producidos o recibidos por una persona o institución durante el curso de su gestión o actividad para el cumplimiento de sus fines y conservados como prueba e información” (Heredia, 1991, p.123); de este modo, la procedencia del documento, en un contexto institucional o para un fin específico, es diferenciador de cualquier documento en general. En el caso de México para la Ley General de Archivos de 2018 en su artículo 4o, fracción XIV, cuando se habla de documento de archivo se refiere:

A aquel que registra un hecho, acto administrativo, jurídico, fiscal o contable producido, recibido y utilizado en el ejercicio de las facultades, competencias o funciones de los sujetos obligados, con independencia de su soporte documental” (Cámara de Diputados, 2018, p.4).

La Ley General de Archivos tiene una perspectiva referida únicamente a los servidores públicos que laboran dentro de una institución. Una acepción que ofrece un equilibrio entre la procedencia del documento y el apartado legislativo, nos la proporciona Mundet:

Entidad de información de carácter único, producida o recibida en la iniciación, desarrollo o finalización de una actividad; cuyo contenido estructurado y contextualizado se presenta como evidencia y soporte de las acciones, decisiones y funciones propias de las organizaciones y de las personas físicas y jurídicas (Mundet, 2011, p. 30).

Así pues, Mundet incorpora en su concepción el carácter único de los documentos y su contenido, generando así una conceptualización de documento precisa y contextual para las instituciones públicas, principalmente. A través de las delimitaciones anteriores se reconoce que el documento representa aquellas actividades, actos, decisiones u otros, en un contexto regularmente institucional, el cual se registra en un soporte. De este modo, hacemos notar que el enfoque archivístico atañe a cualquier clase de institución en el desarrollo de sus funciones.

Hasta ahora se ha dado contexto al archivo desde un aspecto gubernamental, pero también los documentos de archivo son producto de las actividades de personas en un contexto académico, artístico o político, por mencionar algunos. Un caso particular es el de los archivos personales, que Gallegos presenta como “aquellos que normalmente integran documentos de carácter personal y documentos relativos a la función desempeñada por el individuo” (Gallegos, 1993, p.17). De forma más específica Ríos menciona que “El archivo personal es aquel que contiene los documentos generados y recibidos por un individuo a lo largo de su existencia, incluyendo todas sus funciones y actividades, independientemente del soporte” (Ríos, 2008, p.52).

Finalmente, la cualidad fundamental de un archivo personal la comenta del Campo Puerta, al señalar que:

Es la reunión de documentos derivados de la vivencia y/o actividad de una persona o entidad, ganando en valor e importancia cuando traspasa los límites del anonimato y es expuesto para el resto de la

sociedad. Si un archivo personal no sale a la luz y queda reducido al uso y consulta exclusiva del propio productor, su valor social queda reducido a nada (del Campo, 2020).

Así, los archivos personales tienen un gran alcance y relevancia, por lo que deben ser considerados como fuente de información, sin necesariamente estar ligados desde un inicio al contexto gubernamental. Cortés lo define de forma integral y sintética al decir que es:

El conjunto de documentos acumulados en un proceso natural por una persona o institución, pública o privada, en el transcurso de la gestión de asuntos de cualquier índole, los producidos y los recibidos en cualquier fecha, los cuales se conservan y custodian para servir de referencia, como testimonio o información para las personas responsables de tales asuntos y sus sucesores (Cortés, 1982).

El archivo personal, sirve entonces como una fuente de información en forma de un conjunto documental diverso, cuya trascendencia está ligada al alcance o impacto social que una persona o familia tuvo en el transcurso del tiempo.

Por último, hay otra clase de archivos denominados privados, los cuales son producto de las actividades empresariales. El Diccionario de Terminología Archivística¹⁰ enuncia así al archivo de empresa: “Es el archivo que reúne los fondos documentales producidos por una organización mercantil o industrial, de carácter público, mixto o privado”. Otra aproximación a considerar para profundizar es la siguiente:

Es el conjunto de documentos recibidos y generados en el desarrollo de sus funciones por las sociedades mercantiles, necesarios para el cumplimiento de sus objetivos de producir y comercializar bienes y servicios, para la defensa de sus intereses y utilizados también para la configuración de una imagen de marca corporativa (González Pedraza, 2010).

En los documentos está plasmada la información de operación de la entidad privada, la cual es un activo fundamental en la gestión de la organización, ya que su acceso permite una buena toma de decisiones. La información que se resguarde en el archivo, según se gestione, podrá convertirse en conocimiento al permitir identificar estrategias nuevas de

¹⁰ Diccionario de Terminología Archivística. Madrid: Subdirección General de los Archivos. Disponible en <https://www.culturaydeporte.gob.es/cultura/areas/archivos/mc/dta/diccionario.html> (Consultado el 19 de enero de 2023)

acción. Por último, los archivos de empresa son a su vez -como en el caso de cualquier organización- una evidencia de su evolución con el paso del tiempo, que se convertirá en su memoria y le brindará su identidad en comparación con otra.

En el caso de los archivos empresariales, Canto Collí (2017) insiste en que la información, al ser un recurso estratégico para la organización, conlleva a que se produzcan sistemas para su gestión, difusión al interior de la intervengainstitución, almacenamiento y disposición. Luego, corresponde al archivista valorar la información que se produce dentro de cualquier organización y establecer procesos de gestión documental que eviten el riesgo de pérdida de información.

[El documento de archivo digital y electrónico](#)

En la actualidad, las tecnologías de la información incorporan cambios en la práctica archivística; entre ellos, el que más destaca en su alcance es aquel referido a la producción del documento de forma digital o la digitalización del documento.

En México entre los investigadores más destacados con relación al documento digital se encuentran Alicia Barnard y Juan Voutssas, quienes tienen una larga trayectoria en la investigación archivística. Para delimitar al documento digital haré uso de su conceptualización de documento de archivo digital. Primero, nos dicen que “Un documento de archivo es un documento producido (elaborado o recibido, y guardado –separado, apartado–para acción posterior y referencia) por una persona física o jurídica en el curso de una actividad práctica como instrumento o subproducto de tal actividad” (Barnard y Voutssas, 2014, p.80); esto nos indica que el documento en primera instancia debe ser producto de una entidad en el ejercicio de sus funciones. Posteriormente, la definición termina de este modo: “Cuando el documento de archivo es guardado y utilizado en forma digital entonces es un documento de archivo digital, independientemente de la forma original en la que haya sido elaborado y recibido” (Barnard y Voutssas, 2014, p.80). De modo que la procedencia del documento es elemento clave para sustentar su condición como condición de documento de archivo, con independencia de la tecnología utilizada para su producción.

Cabe destacar otro término utilizado en la terminología archivística, conocido como documento electrónico y que se define por los mismos autores como “un mensaje elaborado con base en impulsos electrónicos y que puede ser almacenado en un dispositivo al efecto, transmitido a través de un canal o red y reconstruido a lenguaje natural o forma original por medio de un equipo electrónico, el cual es indispensable para poder hacer inteligible el documento de nuevo” (Barnard y Voutssas, 2014, p. 83-84). Esta caracterización resulta más amplia, ya que abarca documentos como: CDs, DVDs, discos de vinilo, casetes de video VHS, entre otra clase de documentos, ya que todos estos necesitan un aparato electrónico para ser reproducidos y así conocer su contenido.

En el marco de esta conceptualización es necesario señalar lo que en la archivística es conocido como documento nacido digital. Los Archivos Nacionales de Reino Unido nos brindan la siguiente enunciación: “El documento nacido digital es aquel que han sido creado de forma nativa en formato digital (en lugar de ser digitalizados a partir de documentos de archivo en papel)” (The National Archives, s.f.). El ejemplo más común son aquellos documentos creados por un procesador de textos, que a la vez tienen la condicionante de ser visualizados mediante el uso de software. Otras aproximaciones sobre los documentos nacidos digitales la brindan Los Archivos Nacionales de Australia: “son unidades discretas de información digital que se pueden leer, escribir y copiar” (The National Archives, s.f.). En síntesis, el documento nacido digital es aquel hallado en el medio digital generado a partir del software y hardware que sirven para interactuar con él.

Particularmente se centrará la atención en el documento nacido digital, que progresivamente ha tomado relevancia desde hace algunas décadas debido a la proliferación en el uso de las tecnologías de la información para su creación y transferencia. Por solo mencionar un ejemplo de lo anterior podemos considerar el correo electrónico, que involucra el uso de computadoras para la creación, envío y almacenamiento de documentos digitales.

Durante los últimos años la tecnología, en particular aquella referida a la computación, ha visto una gran transformación e implementación en lo referente a los sistemas inteligentes. Parte de este hecho viene derivado del incremento en poder computacional, lo que vuelve viable la implementación en entornos privados y públicos. Paralelamente a los estudios dentro de la IA y el archivo, poco a poco se ha tornado hacia la atención académica a partir del soporte digital. El documento digital o aquel que ha sido digitalizado, brinda más posibilidades de tratamiento en la información. Realizar una aproximación a la definición de IA no resulta sencilla ya que no hay un consenso acerca para definirla de manera única. Una primera definición sobre la IA, como característica de un sistema computacional, es la que nos brindan Kaplan y Haenlein:

La capacidad de un sistema para interpretar correctamente datos externos, para aprender de dichos datos y emplear esos conocimientos para lograr tareas y metas concretas a través de la adaptación flexible (2019, p.17).

Aquí se plantea la IA como una alternativa para diseñar un programa computacional que realice una cierta actividad, en la cual viene propuesto un componente de aprendizaje - que habría que desglosar a mayor detalle- y una valoración de la efectividad. En atención a esta conceptualización, es importante no caer en el error de pensar que el objetivo último de la IA es reemplazar al ser humano en aquellas actividades donde se alcance un mismo nivel de eficiencia o incluso mayor.

La IA debería ser vista como un complemento para realizar las diferentes actividades en las que ésta intervenga. El ser humano será quien gestione la IA y siempre velará por su correcto uso, del mismo modo que otras tecnologías que se han integrado a sus labores cotidianas tales como: las computadoras, el internet, los programas informáticos, entre otros. Las tecnologías anteriores han automatizado algunas actividades o las han vuelto más eficientes y se han integrado a las actividades como otro elemento más de un proceso determinado.

Dentro de la IA se encuentra el campo de conocimiento llamado Aprendizaje Automático (AA), donde la idea de aprendizaje se delimita con mayor precisión. Según Murphy, se define de la siguiente manera:

El conjunto de métodos automatizados que de forma automática detectan patrones en los datos y utilizan los patrones descubiertos para predecir datos futuros o para realizar otro tipo de toma de decisiones bajo incertidumbre. (Murphy, 2012, p.1)

El funcionamiento general descrito por Murphy acerca del AA establece el posible uso que la IA del tipo AA puede aplicarse a los documentos digitales de archivo en lo que respecta al procesamiento de la información contenida en ellos, ya que la información se compone de datos. El AA puede proporcionar así asistencia inteligente cuyo funcionamiento se basa en los datos suministrados y en los algoritmos correspondientes para realizar un aprendizaje determinado sobre ellos.

Hay varios tipos de AA que, según Russell & Norvig (2009), pueden clasificarse en los siguientes:

- **Aprendizaje supervisado:** identifica ejemplos de entrada-salida y aprende una función que asigna la entrada a la salida; todo ello a partir del suministro de casos resueltos; por ejemplo, una serie de imágenes descritas o etiquetadas donde se indica el contenido: “esto es un perro”, “esto es un auto”, “esto es un rostro”. A partir del entrenamiento con ese conjunto etiquetado, el modelo adquiere la capacidad de identificar objetos en imágenes que analiza por primera vez.
- **Aprendizaje no supervisado:** aprende patrones de los datos de entrada, aunque no se le proporcione información explícita; es decir, encuentra un conjunto de características en común en aquella información que se le suministra. El patrón más común es la agrupación (*clustering*): detectar grupos potencialmente útiles de ejemplos de entrada. Por ejemplo, con base en datos del tráfico diario, sería posible hallar días que siguen patrones similares, sin que se le programe una caracterización de forma explícita, utilizando datos asociados con cada día.

- **Aprendizaje por refuerzo:** el agente computacional aprende a partir de una serie de recompensas o castigos. Por ejemplo, en una partida de ajedrez si la IA resulta victoriosa, a través de esquema se buscará ponderar cuáles de las acciones anteriores al refuerzo fueron relevantes para su acierto.

Cabe destacar otra variante del AA, la cual es denominada Aprendizaje Profundo (AP). Skansi la presenta como:

Un tipo especial de aprendizaje con redes neuronales artificiales profundas, aunque hoy en día el aprendizaje profundo y las redes neuronales artificiales se consideran el mismo campo. Las redes neuronales artificiales son un subcampo del aprendizaje automático, que a su vez es un subcampo tanto de la estadística como de la inteligencia artificial (IA)” (Skansi, 2018, p. V).¹¹

Por lo que el AP se puede considerar una aplicación más específica y avanzada, que tiene una particularidad referida al uso de redes neuronales.

Retomaremos la caracterización que propone Schmidhuber sobre las redes neuronales:

Una red neuronal (RN) estándar está formada por muchos procesadores simples y conectados llamados neuronas, cada uno de los cuales produce una secuencia de activaciones de valor real. Las neuronas de entrada se activan a través de los sensores que perciben el entorno, otras neuronas se activan a través de los diferentes pesos de las neuronas previamente activas. (Schmidhuber, 2015, p. 4)

A grandes rasgos, el funcionamiento de una red neuronal toma inspiración en las redes neuronales que se hallan en el cerebro humano, organizándose en un entramado de capas que pueden tener funciones especializadas. Cada neurona tiene un valor o también llamado peso determinado, hay diferentes capas de neuronas que mediante la función de

¹¹ No todas las definiciones de AA lo demarcarían como un campo de la Estadística, puesto que hay otras herramientas matemáticas que se implementan de forma computacional y se reconocen dentro de este campo. Es posible que el autor lo enuncie de esta forma para enfatizar la relevancia del estudio matemático de los datos en el AA. Además, sí hay una aceptación en múltiples fuentes para ubicar el AA como un subcampo de la IA.

activación se les asignan distintos valores, para la resolución del problema. Así, para cada problema particular se busca una configuración particular de neuronas que establezca los valores o pesos que lo resuelvan, refinándose a través de un periodo de entrenamiento. El AP resulta ser una forma muy potente de IA, que permite tratar de resolver problemas con un mayor grado de complejidad.

El estudio de la IA y sus variantes implica un conocimiento especializado por parte de los expertos en el área, pese a ello, un acercamiento conceptual es necesario para entender cómo estas nuevas tecnologías transforman la manera de procesar la información. En términos generales la IA ha conseguido avances importantes en cuanto a la automatización de diversas tareas, lo que motiva a su aplicación en documentos de archivo sea cada vez más frecuente, promoviendo a su vez un interés académico hacia su pertinencia y sus usos responsables.

Dentro del área de la IA y las ciencias de la computación se encuentra el campo conocido como Procesamiento de Lenguaje Natural (PLN) el cual es descrito por Reshamwala et al. (2013), como una rama de las ciencias de la computación, la inteligencia artificial y la lingüística preocupada con las interacciones entre los ordenadores y el lenguaje humano. Esta área en particular es en especial relevante para los documentos de texto, hallados en el archivo, y por ende el uso de esta tecnología puede ser beneficiosa para acelerar la gestión de los documentos. En particular el PLN aborda varias tareas, algunas de las cuales pueden resultar relevantes para la archivística, por ejemplo:

- **Reconocimiento de entidades:** Identificar palabras o frases concretas ("entidades") y clasificarlas, por ejemplo, como personas, lugares, enfermedades, genes o medicamentos (Nadkarni et al., 2011)
- **Clasificación de textos:** Proceso de clasificar documentos en un conjunto predefinido de categorías en función de su contenido. Esta asignación puede utilizarse con fines de clasificación, filtrado y recuperación (Mohammad et al., 2016).

- **Modelado de tópicos:** Una forma de examinar temas en grandes colecciones de documentos, que considera que cada documento es una mezcla de temas latentes que puede ser identificada para elaborar una descripción (Martin & Johnson, 2015)

Las tareas anteriores pueden ayudar en términos de la organización de documentos, tanto en su clasificación, como en generar una primera versión de descripción, así como en la búsqueda de información concreta.

Sesgos Algorítmicos

Friedman y Nissenbaum entienden el sesgo en el contexto de la computación como: “los sistemas informáticos que discriminan sistemática e injustamente a determinadas personas o grupos de individuos en favor de otros” (Friedman & Nissenbaum, 1996, p. 332). El sesgo fácilmente puede poner en riesgo de sufrir discriminación a las personas cuando está presente en un algoritmo que se utiliza de forma irreflexiva. Reconocer los tipos de sesgos es un primer paso para mitigar su aparición. Friedman y Nissenbaum (1996 pp.336-336) identifican tres tipos de sesgos:

- **Sesgo Preexistente:** Tiene sus raíces en instituciones, prácticas y actitudes sociales que pueden reflejar los prejuicios personales que interfieren en el diseño del sistema, esto puede ocurrir de forma explícita o inconscientemente durante el desarrollo. Algunos ejemplos son la discriminación racial o por género.
- **Sesgo Técnico:** Se refiere a una resolución incorrecta de problemas planteados que involucran tanto el hardware como el software, así como los periféricos utilizados. Principalmente se relaciona con el modelado de constructos sociales fuera de contexto, o bien, mediante la cuantificación errónea de elementos cualitativos.
- **Sesgo Emergente:** Aparecen durante el uso del sistema, esto debido a condiciones inesperadas o a un contexto social o cultural distinto.

Dado que los algoritmos son un reflejo de sus desarrolladores, detectar sesgos preexistentes requiere un ejercicio de reflexión crítica desde que se concibe la propuesta y durante todo el proceso. En cuanto a los sesgos técnicos, es conveniente contar con un equipo interdisciplinario, de modo que profesionales especialistas en los datos puedan brindar luz para reconocer un problema. Prevenir los sesgos emergentes requiere una vigilancia sobre el comportamiento del algoritmo, para identificar posibles escenarios o eventos no anticipados durante el desarrollo.

Danks y London proponen otra clasificación de sesgos algorítmicos, profundamente relacionada con el proceso de desarrollo de la IA (Danks & London, 2017, pp. 2-4):

- **Sesgo en el entrenamiento de datos:** Causado por desviaciones en los datos de entrenamiento o de entrada proporcionados al algoritmo; por ejemplo, tener más datos de una sola localidad cuando la IA se utilizará a nivel nacional.
- **Sesgo de concentración:** Se refiere al uso o no de ciertos datos en el entrenamiento del algoritmo, ya que esto lo podría sesgar. Un ejemplo sucede en la decisión de tomar en cuenta el género, la raza, el código postal o el lugar previo de estudios como variable para considerar la admisión a una universidad.
- **Sesgo de procesamiento:** Siendo conscientes que los datos de entrenamiento pueden estar sesgados, una forma de mitigarlo es mediante el uso de algoritmos que permitan suavizar o normalizar los datos; sin embargo, esto puede agregar otros desbalances, de los cuales también se deberá ser consciente.
- **Sesgo del contexto transferido:** Se produce cuando se utiliza un algoritmo fuera de los contextos previstos, que puede no funcionar de acuerdo con lo esperado. Aquí podría tomarse el caso de un vehículo autónomo que es entrenado con datos de una cierta localidad, pero al utilizarlo en otra infringe reglas de tránsito que no habían aparecido en su entrenamiento.

Las clasificaciones anteriores de sesgos algorítmicos, aunque no sean exhaustivas, brindan luz acerca de consideraciones que darán confiabilidad y evitarán múltiples

iteraciones futuras en el desarrollo de una IA. Para lograr mitigar algunos de los sesgos algorítmicos aquí expuestos, Danks y London (2017, pp. 4-6) proponen:

- 1) **Identificar el sesgo:** La comprensión del rol que tendrá la IA en el contexto social donde se implementa, así como de las normas éticas y jurídicas básicas que son relevantes, permitirá anticipar sus consecuencias y eliminar efectos negativos.
- 2) **Intervenir el sesgo:** Para mejorar el algoritmo, es posible restringir su operación, rediseñarlo, diseñar uno nuevo o reconsiderar la ponderación asignada a sus resultados durante la toma de decisiones, proporcional a la dificultad de mitigar o eliminar el sesgo.

Es necesario tener conciencia de la dificultad de producir un algoritmo sin sesgos, por lo que deben realizarse acciones para anticipar y evitar sus efectos durante todo el proceso de desarrollo. Otras formas de mitigar el sesgo propuestas por Lee et al. (2022) son:

- **Evaluación del impacto algorítmico:** Como práctica de autorregulación, durante las fases de diseño, implementación y supervisión se pueden realizar y documentar indagaciones sobre el impacto de sus decisiones, que permitan sondear y evitar cualquier sesgo potencial que esté incorporado en la decisión algorítmica o que sea resultado de ella.
- **Equipos de trabajo interfuncionales y expertos:** Reunir a expertos de varios departamentos, disciplinas y sectores ayudará a facilitar normas de responsabilidad y estrategias para mitigar los sesgos, antes y durante la implantación del modelo.
- **Auditoría regular del algoritmo:** Una auditoría periódica para revisar los datos de entrada y las decisiones de salida, junto con la información recabada de los desarrolladores, la sociedad civil y otras personas afectadas por el algoritmo, permitirá detectar mejor los sesgos y, posiblemente, disuadirlos.
- **Involucramiento del factor humano en la mejora:** Continuar con la identificación y corrección de resultados sesgados después de que se desarrolle, pruebe y

ponga en marcha un algoritmo, así como la documentación de los sesgos hallados, es una buena práctica que se debe asumir de forma permanente.

Como una práctica permanente debe asumirse la presencia de sesgos, para no asociar a la IA como un mecanismo totalmente objetivo y sin errores. Se debe entonces ser consciente del alcance y limitaciones que tiene una IA, comprender el grado de ayuda en la toma de decisiones que puede lograr y proponer esquemas en donde estas decisiones puedan ser supervisadas por un humano o ser reversibles. Para Danks y London (2017, p. 6) los involucrados en el diseño y los responsables de su uso deben colaborar para ello de acuerdo con las mejores prácticas, en beneficio de los usuarios del sistema; así, el espacio de creación ofrece, tanto la posibilidad de introducir sesgo, como la de compensarlo.

El panorama profesional del archivista en el uso de la inteligencia artificial en el archivo

El abordaje de la IA en el archivo

Explicabilidad e interpretabilidad de la IA en tratamiento de la información

Un primer acercamiento a la IA en el archivo parte de entender la forma en que nosotros, como archivistas, nos posicionamos con respecto a ella; dicho de otro modo, establecer qué clase de acciones es posible tomar al estar en contacto con esta tecnología en expansión.

Partir desde el aspecto teórico puede ser un buen inicio. Andresen (2019, p.130) realiza un acercamiento a la IA a la luz de la producción del documento, donde señala que los algoritmos, a diferencia de los humanos, no pueden rendir cuenta del proceso de registro de la información contenida en el documento. Resulta clave identificar un cambio de perspectiva siendo que, en contraste con lo que sucede con las personas, la IA por sí misma no podría rendir cuenta del proceso que realiza, sino que son las personas que la desarrollan junto con las encargadas de utilizarla quienes tienen la responsabilidad de hacerlo.

Por parte de la teoría archivística, Andresen (2019, p.131) cita la reflexión de Schellenberg donde los documentos son productos de distintas actividades de cada organización; aquellas políticas o procedimientos que controlan a estas actividades son de importancia al ayudar a entender la información de ellos. Derivado de lo anterior, establece como aportaciones desde la teoría de la gestión documental los documentos sobre políticas y procedimientos.

El enlace de las políticas sobre los documentos, las cuales aplican regularmente a personas, se traslada en la actualidad hacia analizar como ahora la IA gestiona la información y comprender la razón de un resultado determinado en el tratamiento de la

información. Al ser archivistas podemos establecer una relación con la IA a partir de entenderla como factor tecnológico implicado en la producción de documentos, sobre el cual deberían desarrollarse políticas o procedimientos que gestionen la IA.

Andresen (2019, p.132-134) señala que se debe entonces documentar el funcionamiento del algoritmo, y lo más sencillo sería establecer una explicación sobre su operatividad. Propone entonces dos posibles estrategias, siendo la primera de ellas explicitar por qué y cómo el algoritmo dio una determinada solución, con la desventaja de conducir a un rango de explicaciones que podría ser muy extenso. Este autor utiliza los conceptos de “explicación” y “predicción”: una explicación refiere al cómo y al por qué, mientras que la predicción tiene que ver con aproximar la justificación de un resultado con base en tratar de anticipar el probable resultado. De esta forma, se podría entender a la predicción como una simplificación del entendimiento, mientras que la clarificación requiere que el conocimiento profundo del algoritmo sea accesible para quien la revisa.

En el caso del archivista, que requerirá valorar el alcance e impacto de la IA, es recomendable que en la medida de lo posible logre acceder a una fundamentación a grandes rasgos sobre cómo funciona la IA. Investigar por ejemplo cómo un determinado tratamiento sobre los documentos, donde la accesibilidad implique también que pueda interpretarla desde su conocimiento especializado, cuidando que dicha explicación o predicción no deje fuera criterios importantes. En la medida en que, tanto los responsables del uso de la IA como los usuarios comprendan en términos razonables para su propia disciplina, el uso de estas tecnologías dejará que causar tanta incertidumbre y podrá generar confianza en los beneficios que tiene.

Los algoritmos de aprendizaje automático hacen inferencias con base en los datos que les son suministrados, así que un modo por el cual es posible tratar de inferir una justificación sería mediante un ambiente donde haya un control de los datos. Esto a su vez habilitaría realizar una evaluación de los riesgos o, en su caso, una posible evaluación ética del algoritmo. Si bien el archivista no posee los conocimientos técnicos, si puede participar

en este ejercicio de valoración a partir del entendimiento de los documentos de entrada, el análisis de los resultados, los criterios enunciados en la explicación y la consecuencia de cada decisión.

Finalmente, Andresen (2019, pp.136-140) señala la importancia de investigar qué clase de documentos procedentes de los procesos, explicaciones y predicciones pueden servir para documentar al algoritmo y dar cuenta de su funcionamiento. Esto se alinea con la transparencia y el derecho a la explicación en regulaciones como el Reglamento General de Protección de Datos (GDPR).

Las regulaciones en materia de protección de datos deberían tener en cuenta evaluaciones éticas o de riesgo sobre los algoritmos. Si el caso fuera que una empresa privada brindara el servicio de la IA, sería recomendable vigilar que evalúe los algoritmos a aplicar en estos términos, cómo parte de sus funciones. Sería conveniente documentar el funcionamiento del algoritmo, junto con una política de control de información. Desde la gestión documental sería deseable abordar específicamente cómo se gestionan los documentos que están ligados a la inteligencia artificial.

En un contexto donde múltiples documentos serán resultado del uso de la IA, la explicabilidad resulta especialmente relevante al tratar de identificar la fiabilidad del documento, generado con la intervención de una IA. Aquí es conveniente insistir en que una IA también puede participar en un proceso de decisión que posteriormente implique la creación de un documento, quizá de forma no automatizada. Desde esta perspectiva, la IA puede formar parte de su producción y, si es el caso, desde una capa conceptual, por lo que conviene que sea detallado su uso. De manera similar, si una IA fue parte del proceso de organización, valoración u otro capaz de ser aplicado por esta en el documento, es crucial que dentro de los metadatos que acompañan a los documentos se indique cual fue su uso cuándo fue empleada.

Como hemos visto, tres ejes fundamentales de la discusión sobre la IA en el archivo son la transparencia, la regulación y la ética, por lo que la responsabilidad aparece como un

aspecto crucial. La disrupción que la IA provoca en la gestión documental se debe en gran parte a que la cadena de gestión de los documentos ya no es realizada únicamente por seres humanos, sino también por la IA (Bunn, 2020, p. 145). Entonces, ¿Cómo serán las pruebas de esas transacciones? En un sistema de gestión de documentos, por ejemplo, cabe la posibilidad del uso de la IA, tal como se vio anteriormente. Cuando se utilice una IA siempre será conveniente que se integren los metadatos para reconocer que participación fue realizada por la IA o por el archivista.

Algunos profesionales de la gestión documental ya se enfrentan a la tarea de aplicar el GDPR y sus disposiciones en relación con la toma de decisiones individuales automatizadas. Al abordar esta tarea, quienes se dedican a la gestión documental deben asegurarse de formar parte de las conversaciones sociales e interdisciplinarias más amplias que están empezando a producirse (Bunn, 2020, p.146). En esto, será crucial que el factor ético impulse la regulación de los usos de la IA en plena conciencia de sus alcances, pero también de sus límites.

La interpretabilidad o la comprensión de lo que ha hecho el modelo o de cómo funciona el sistema hacen que la Inteligencia Artificial Explicable (XAI por sus siglas en inglés) sea una actividad altamente técnica, que busca establecer el rigor y la comprensión teórica en el desarrollo de nuevas técnicas de IA¹². De este modo según Bunn (2020, p.145) la XAI se sitúa firmemente en el ámbito más amplio de la investigación en IA como un subcampo en crecimiento. Es un acercamiento nuevo y más interdisciplinario, que empieza a superar los límites del campo de la IA -más estrechamente técnico- y a recurrir a otros campos como la filosofía y las ciencias sociales.

Bunn (2020, p.146), en colaboración con el personal de los Archivos Nacionales y un investigador del campo de la Interacción Humano-Computadora, organizó un taller en el que se pudo explorar el tema de la XAI junto con otros participantes de diversos ámbitos lo cual

¹² El campo de investigación XAI reúne los enfoques que antes fueron presentados por Andresen como explicar y predecir.

dio lugar a un programa de investigación en la interacción con la IA en la institución. Este caso muestra que la consideración del uso de la IA en cada contexto puede contemplar una discusión previa por parte de los implicados, de modo que se comprendan los alcances particulares y las medidas necesarias para comprenderlos. A través de este ejemplo veremos que un taller interdisciplinario puede ser un buen inicio.

Durante la investigación, Bunn (2020, p.147) halló que los participantes analizaron la XAI como una extensión de la necesidad humana de justificación y comprensión. El análisis del funcionamiento se consideraba un proceso iterativo e interactivo, así como un comportamiento humano contextual que contribuye a construir la cohesión social y la confianza. Se determinó que comprometerse con la XAI significaba implicarse también con preguntas como: ¿Cuándo hay que ofrecer una explicación? ¿Cuán detalladas deben ser? ¿Qué es una explicación suficientemente buena en cualquier circunstancia?.

Desde la gestión documental, la transparencia y la responsabilidad son dos de los mecanismos que pueden impulsar la explicabilidad. (Bunn, 2020, p.148) Como se ha dicho anteriormente el hecho de documentar el desarrollo de una IA, al menos si bien no describe profundamente su funcionamiento intrínseco, puede proporcionar elementos relevantes sobre su proceso que sirvan como puntos de partida para la discusión abierta entre los responsables y los interesados.

Podría parecer complicado que siendo archivistas tengamos un conocimiento que nos permita darle una interpretabilidad de la IA dada su complejidad técnica, pero un buen inicio sería entender los conceptos de explicabilidad e interpretabilidad para formar parte de los diálogos con respecto a la IA. Así nos hallamos situados para identificar los documentos que son producto del desarrollo de la IA y establecer su conservación permanente (Bunn, 2020, pp.148-149). En la interacción resulta recomendable partir de los desarrollos teóricos de la XAI y comprender también el papel que tenemos como custodios de los documentos que forman parte del desarrollo de la IA y que pueden brindar luz acerca de su entendimiento.

Bunn (2020) señala que al adentrarnos en el espacio más específico de la XAI y la responsabilidad algorítmica, ya se están conceptualizando algunas nuevas necesidades como registrar en documentos el desarrollo de la IA. Por ejemplo, la declaración de la Association of Computing Machinery sobre la responsabilidad de los algoritmos habla de cómo "debe mantenerse una descripción de la forma en que se recogieron los datos de entrenamiento". A la vez, también el Informe Europeo sobre el Marco de Gobernanza, comienza a dar cuerpo a un "requisito de hojas de datos" en forma de "un documento semiestructurado que hace preguntas como "¿Porqué se creó el conjunto de datos?", "¿Cómo se recogieron los datos?". Además, propone un nuevo documento llamado evaluación del impacto algorítmico, así como un archivo de decisiones de sistemas (Panel para el Futuro de la Ciencia y la Tecnología, 2019).

La gobernanza o regulación de la IA requerirá cierta clase de documentos y es en ellos en los archivistas pueden evitar que se reinvente la rueda y para garantizar que esa nueva clase de documentos creada finalmente sirvan para su propósito (Bunn, 2020, p.150). Así, contribuimos a la discusión de la creación de nuevos documentos donde se utilice la IA, a partir de la documentación en el proceso de creación e involucramiento necesario, teniendo en cuenta que es un trabajo colaborativo.

Bunn (2020, p.151) concluye considerando que los archivistas pueden aplicar inmediatamente su experiencia en documentación para ayudar en estos esfuerzos, abordando (si no lo están haciendo ya) cuestiones como

- ¿Qué documentos se crean en los equipos de investigación de IA para documentar su proceso?
- ¿Qué documentos se crean sobre las decisiones de adquisición o despliegue de sistemas que utilizan IA?
- ¿Qué documentos se crean sobre las decisiones y el impacto de dichos sistemas?
- ¿Son los documentos creados suficientes para cumplir las disposiciones legales

vigentes?

- ¿Los documentos creados cumplen las normas de calidad requeridas?

Así en la documentación del desarrollo de la IA, desde su diseño y desarrollo hasta su implementación, es posible organizar un registro pertinente de la IA, donde además se identifiquen los documentos que pueden brindar transparencia a la par que es supervisada.

El contacto del archivista con la IA

Una visión integradora desde las ciencias de la computación en relación con la IA en el marco de la archivística nos la ofrecen la revisión de literatura realizada por Colavizza et al. (2022, pp.9-11), que resalta que las metodologías generadas por la archivística pueden contribuir a que los datos para el entrenamiento de los algoritmos de IA se implementen en un marco de transparencia. A su vez, este procesamiento automatizado puede formar nuevas clases de archivos en el sentido de cómo son creados y consultados. El contacto del archivista con esta tecnología abre un panorama de discusión sobre su aprovechamiento para la asistencia en las actividades que se pueden llevar a cabo en el archivo.

Hay técnicas de PLN que pueden automatizar actividades con los documentos, como la protección de datos personales y la búsqueda, entre otras. Esto tiene la principal ventaja de acelerar el procesamiento de grandes volúmenes de información, pero existe la preocupación de que tan fiable y exacto es el tratamiento tiene sobre el documento.

Algunas colecciones dentro de los archivos, que podrían estar disponibles para consulta, terminan siendo de acceso restringido debido a las preocupaciones de protección de datos personales (Jailliant y Caputo, 2022, p.826). Una razón por la que estos suceden es que, al contener información privada y datos sensibles, estos documentos requieren un análisis más exhaustivo para evitar posibles vulneraciones a la privacidad y esto termina restringiendo el acceso a ellos como una medida de seguridad.

El uso de herramientas de IA puede ayudar a realizar los documentos en búsqueda de contenido sensible, aunque debe reconocerse que están sujetas a posibles errores. (Jailliant y Caputo, 2022, p. 827).

Los archivistas y académicos tienen que involucrarse en el uso de esta tecnología para entender cómo y por qué se han seleccionado determinados documentos dentro de un gran conjunto de datos. El aprendizaje automático no requiere necesariamente una formación avanzada en informática y esto incluye a aquellos que no tienen experiencia en programación. Es posible crear guías para la utilización de estas herramientas como las que promueve el sitio web *Data Foundry*¹³ (Jailliant y Caputo, 2022, p.830).

La experiencia práctica que el archivista podría tener de la utilización de software de aprendizaje automático, mediante acercamientos guiados, podría generar un conocimiento práctico para así dimensionar su implementación. Steve Rigden, archivista digital de la Biblioteca Nacional de Escocia, identifica una serie de tareas a las que el archivista puede sumarse en el desarrollo del aprendizaje automático (Jailliant y Caputo, 2022, p.831):

- Identificar conjuntos de datos
- Determinar los algoritmos más eficaces para aplicarlos
- Probar y refinar los modelos para los datos de prueba, para dar a la máquina lo que necesita para aprender, y volver a refinar y probar.

Así, no tenemos que ser expertos en los aspectos técnicos de la IA, pero sí tener interés en participar en el desarrollo de estas herramientas como asesores y usuarios de prueba. Puede que al principio halla cierta resistencia hacia al uso de estas herramientas, pues alteran las prácticas ya establecidas en distintos procesos archivísticos, es factible que esto se atenúe al indagar los beneficios que a largo plazo sean capaces de brindar; esta disposición se modificaría al participar en la vigilancia del proceso de su desarrollo y de las fases de pruebas.

¹³ Para más información consultar: <https://data.nls.uk/>

En este panorama donde formamos parte del proceso de diseño en cuanto lo documenta y vigila, cabría profundizar en otros modos concretos en que los archivistas podemos sumarnos a la discusión de la IA: la definición de lineamientos para producir documentos sobre el contexto de los conjuntos de información que se utilizan para producirlas.

Según Gebru un factor fundamental de la evaluación del uso de la IA en los archivos tiene que ver entonces con los conjuntos de datos usados para entrenarlas. En particular, el AA tiene limitaciones al evitar tener sesgos, lo cual puede verse originado con el modo en que los conjuntos de datos son recolectados, anotados y utilizados (Gebru, 2020, p. 2). En el caso específico de la protección de datos personales, el archivista puede estudiar los errores y los efectos no deseados -que pueden producir prácticas de injusticia o discriminación-, para participar en el proceso de rectificarlos.

Hay distintos factores a considerar al preparar un conjunto de datos, siendo uno el cuidado de la privacidad, entre otras cuestiones éticas. Abordar tales factores requiere de explorar otros campos del conocimiento como lo pueden ser los archivos, se ha articulado un dominio con respecto a los procesos de recopilación y anotación de datos. (Gebru, 2020, p.2.) La autora rescata las prácticas de inventario, descripción, organización y la creación de los respectivos instrumentos de descripción archivística (catálogos, inventarios, guías de archivos) y propone cómo pueden ser utilizadas también para aplicarse como formas de control y descripción de datos, al incluir:

- Una declaración de la misión institucional, que defina los conceptos o subgrupos sobre los que se recopilarán datos.
- Definición de responsables de conservación, que deberán sopesar los riesgos y beneficios de la recopilación de diferentes tipos de datos y marcos teóricos para valorar los datos recopilados.

- Códigos de conducta/ética, así como un marco profesional para hacerlos cumplir.
- Formas estandarizadas de documentación.

Para generar e implementar estos instrumentos en la construcción de conjuntos de datos, al igual que para la puesta en práctica de los procesos de gestión documental y conservación de documentos, será invaluable la participación de un profesional especializado.

Los estudios sobre archivos han generado una literatura sobre el tratamiento documental; entre los avances que se han logrado destacan los siguientes: directrices sobre cómo etiquetar los datos; recopilación y accesibilidad de información privada, intercambio de conjuntos de datos entre plataformas, reflexiones críticas sobre la diversidad y la inclusión, teoría de la valoración y la selección. (Gebru, 2020, p.3). Así, el desarrollo que la archivística ha tenido durante los años puede ser llevado a los conjuntos de datos aprovechados para el entrenamiento de la IA. Al realizar la inferencia sobre los documentos como formas en contexto que contienen los datos, es posible contribuir para su correcta gestión.

Gebru realiza la distinción entre los conjuntos de datos utilizados en el aprendizaje automático, los cuales tienen el objetivo de mejorar las métricas de precisión en tareas definidas, y aquellos conjuntos de documentos que se manejan en los estudios sobre archivos, que implican: preservar la herencia y memoria, educar e informar. Sin embargo, propone que, de manera particular, la autenticidad, la privacidad y la inclusión deberían ser elementos involucrados en la supervisión de la generación de los conjuntos de datos en el aprendizaje automático (Gebru, 2020, p. 3). Así, los requisitos que deben cumplir los documentos en un archivo deberían ser necesarios también para los conjuntos de datos de entrenamiento.

La transparencia y la responsabilidad son principios fundamentales de la ética archivística. Para defender estos principios, los archivos se rigen por la gestión documental: en los archivos se guardan documentos de las decisiones y evaluaciones del flujo de valoración, lo cual implica que el proceso pasa por muchos niveles de supervisión por parte

de archivistas, conservadores, productores de documentos y gestores de documentos. Así los conjuntos de documentos, antes de ser llevados a la IA, deberían pasar por varios filtros correspondientes, para asegurar el manejo ético en su recolección y su adecuada documentación, que habilite la transparencia y la comprensión en el futuro funcionamiento de la IA.

Según Gebru (2020) un flujo de valoración documental aplicable a los documentos entendidos como datos o fuentes, de forma breve en la Tabla 6 consiste en lo siguiente:

Proceso del flujo documental	Acciones específicas
1. Declaración de la misión	Nivel más alto de formulación de la agenda que determina los temas/conceptos de interés.
2. Política de desarrollo de la colección	Una política más específica sustentada en la misión sobre lo que se recoge, lo que no, y dónde y cómo buscar las fuentes.
3. Valoración	Evaluación basada en criterios sobre si una determinada selección de fuentes merece ser recogida.
4. Procesamiento/indización (microevaluación)	<p>Procesar las fuentes individualmente o a nivel de carpeta/documento, y la actualización de los instrumentos para su búsqueda</p> <p>Las fuentes pueden ser descartadas por motivos de privacidad o por irrelevancia.</p>

Tabla 6 Propuesta de Flujo de Valoración para datos. A partir de Gebru 2020

De forma específica en lo que refiere a privacidad y a ética, la comunidad de archivistas ha generado varios niveles de códigos de conducta profesional que guían y hacen cumplir las decisiones acerca del tratamiento que tienen los documentos. Las organizaciones que agrupan a los archivos, las bibliotecas y los museos tienen códigos de ética y conducta individuales, como la valoración documental, el acceso a la información sensible y la propiedad intelectual, por mencionar algunos (Gebru, 2020, p.7). Las políticas y procedimientos que ha desarrollado la archivística respectivamente pueden ser tomados como base, en tanto que el archivista puede colaborar en el diseño preciso de políticas y procedimientos para la integración del conjunto de datos.

En primer lugar, Gebru propone una política para establecer un sistema de membresía que refuerce el comportamiento ético al momento de generar un conjunto de datos. En segundo lugar, propone el establecimiento de organizaciones interinstitucionales que estén por encima de los empleadores directos puede ayudar a garantizar que los principios éticos resistan a las motivaciones lucrativas. (Gebru, 2020, p.7). Finalmente el control de los conjuntos de datos, tiene relación directa con la persona que los gestiona y después los utiliza, así que la ética debe ser el eje principal, antes incluso que generar cualquier valoración, ya que es en la medida en que se asume una responsabilidad las políticas, procedimientos, instrumentos y formas de descripción se llevarán a cabo del mejor modo posible

Principios de diseño para un software de procesamiento de documentos

Hutchinson, a partir de una revisión de herramientas de PNL aplicables a documentos para los propósitos de este informe, se dará prioridad a las conclusiones que el autor propone a partir de ella, al establecer un conjunto de principios de diseño que todo software que utilice PNL en el contexto de los archivos debe adoptar (Hutchinson, 2020):

- **Utilizable:** Las herramientas deben diseñarse suponiendo un nivel adecuado de conocimientos técnicos por parte del usuario.

- **Interoperable:** Las herramientas deben tener capacidad para integrar los resultados del procesamiento de PNL y el aprendizaje automático en otros elementos del flujo de trabajo de procesamiento digital de una institución; así como proveer soporte para compartir resultados y modelos entre colecciones e instituciones.
- **Flexible:** Las herramientas deberán tener capacidad para profundizar en el nivel del documento a través de la interfaz de usuario, pero también de exportar los resultados para su procesamiento y análisis independiente, aplicando también herramientas de visualización.
- **Iterativo:** Las herramientas deberán tener la capacidad de entrenar y perfeccionar los modelos, idealmente a través de una interfaz de usuario, identificando los falsos positivos y los elementos omitidos.
- **Configurable:** Las herramientas deberán brindar la capacidad de perfeccionar el modo en que se aplican las técnicas de PNL y aprendizaje automático, como las opciones de limpieza de datos, las listas de palabras que deben omitirse - por ejemplo, aquellas que son muy frecuentes pero que en realidad no aportan información relevante- y los modelos estadísticos utilizados.

Es importante señalar que este conjunto de principios de diseño considera que el uso de este software no se limita únicamente a la realización de la tarea del procesamiento del lenguaje natural, sino también a su capacidad de integración en flujos de trabajo. En la medida en que el software cumpla estos requerimientos, será más sencillo para el archivista ser preciso en su aplicación. Además, las diversas herramientas identificadas pueden ser implementadas sin la necesidad de llevar a cabo un desarrollo desde cero, sino que pueden ser adaptadas y configuradas para su ejecución sencilla y práctica.

Hutchinson concluye al mencionar que las herramientas descritas podrían tener una aplicación más amplia. Por ejemplo, reconocimiento entidades que aparecen en el texto e identificar los temas principales de un documento son funcionalidades que podrían utilizarse tanto para la valoración como para la descripción. También existen oportunidades para

ejecutar estas herramientas en las fichas de descripción de documentos, así como en los documentos de archivo digitalizados y los nacidos en el ámbito digital. Estas herramientas tienen un gran potencial para ser desarrolladas e integradas para un uso más amplio por parte de la comunidad archivística. (Hutchinson, 2020, p. 168).

Casos de estudio en el uso de la IA en los archivos

A continuación, se describen casos relacionados al uso de la IA en archivos donde el profesional tenga un rol activo en el proceso de investigación de evaluación, implementación o uso.

El primer estudio piloto analizó el desarrollo de un software de IA que pudiera asistir en la valoración documental de documentos digitales. De forma específica, el algoritmo de IA de tipo aprendizaje automático tenía como función el procesamiento de los textos para establecer categorías temáticas (Makhlouf Shabou et al., 2020, p.175). Este caso es particularmente interesante, pues muestra que la IA puede auxiliar en el reconocimiento de la información de grandes cantidades de documentos. El establecimiento de categorías temáticas puede ser especialmente útil para identificar los tipos de documentos.

En este proyecto se procesaron documentos con información no estructurada, que son aquellos donde hay campos de información claramente definidos, como sucede en las cartas oficios, correos o memorándums, que regularmente incluyen textos en formato libre. El reto consistía en manejar casos extremos de valoración de documentos en los que el archivista tuviera que gestionar soportes digitales sin ninguna indicación de su naturaleza y su contexto. Para hacer frente a este reto, la idea explorada era utilizar al máximo el enfoque de aprendizaje automático para procesar datos. El objetivo de la investigación era desarrollar una prueba de concepto para una herramienta de valoración de documentos digitales que ayude a tomar decisiones sobre el archivo o la eliminación de datos corporativos (Makhlouf Shabou et al., 2020, pp. 176.).

Los principales usuarios potenciales de esta herramienta de valoración propuesta serían los archivistas, que deben llevar a cabo un proceso de valoración riguroso de forma documentada. Un uso semejante de la IA puede servir para establecer una visión general del archivo incluso antes del inventario, lo que podría permitir un reconocimiento más eficiente de los documentos y podría verse integrado como una nueva etapa en el proceso de organización. Este estudio pretendía explorar los mecanismos y enfoques de valoración de datos, documentos y archivos de forma sistemática y/o automatizada (Makhlouf Shabou et al., 2020, p.177).

Se seleccionaron cinco expertos para realizar la prueba, en la que se evaluaron los conocimientos de los candidatos sobre los procesos, los criterios y las herramientas de valoración de los documentos. Los candidatos tuvieron que leer la documentación disponible sobre las métricas propuestas antes de ser elegidos. También tuvieron la oportunidad de hacer comentarios o sugerencias de mejora. Se realizó otra prueba con el mismo grupo de expertos sobre las métricas y funcionalidades de la minería de textos (Makhlouf Shabou et al., 2020, p.180). Este proceder ilustra que la retroalimentación entre los usuarios finales y los desarrolladores de la IA es esencial para así evaluar la calidad de los procesos asociados al archivo. El desarrollo de IA requiere de un proceso integral, donde se consideren los requerimientos y necesidades de las personas que usarán principalmente el sistema.

Se tenía una colección de documentos llamada dataOAEN, la cual tenía un tamaño de 102 GB y constaba de 57,286 documentos, de los cuales sólo 13,179 habían sido indexados con metadatos y/o texto completo. Los otros documentos estaban en formato de imagen (.jpg o .tiff) y correspondían a escaneos que no habían sido indexados en el sistema. Para estos documentos se propuso en un estudio futuro utilizar la tecnología de reconocimiento óptico de caracteres *OCR (Optical Recognition Character)* en inglés, para extraer la información. La prueba de concepto, desarrollada tenía la funcionalidad de reconocimiento de entidades con nombre (NER) que se probó para el reconocimiento de nombres propios y fechas. Se utilizaron los nombres de las personas que trabajan en la oficina para verificar el enfoque (Makhlouf Shabou et al., 2020, p.181). Se observa aquí que,

en circunstancias normales identificar de forma manual tal cantidad de documentos tomaría bastante tiempo, pero al utilizar un algoritmo como el NER, brinda una optimización drástica que sirve en la aceleración del reconocimiento del contenido en los documentos.

Los conjuntos de aprendizaje y entrenamiento se crearon a partir de una consulta con palabras clave relacionadas con los documentos de las minutas en el contenido y el título. El tema de clasificación utilizado en la prueba de concepto fue la identificación de documentos de actas de reuniones (Makhlouf Shabou et al., 2020, p.181). Aquí cabe destacar que incluso la IA, aunque parezca de uso general y versátil, debe ser entrenada con documentos específicos del archivo en cuestión a tratar.

El uso de la minería de texto demostró que, entre mayor fuera el conjunto de prueba, con más exactitud se realizaba la identificación de los documentos, pudiendo alcanzar hasta un 95.8%; cabe señalar que la medida de precisión calcula el porcentaje de actas correctamente identificadas de todas las posibles (Makhlouf Shabou et al., 2020, p.189).

Este estudio finalmente detalla de manera particular una forma de abordar el uso de aprendizaje automático para documentos de tipo administrativo, el cual tuvo altas tasas de éxito, sin embargo, no se especifica su posterior implementación más allá de ser una prueba de concepto. Es interesante señalar aquí cómo, desde nuestra experiencia profesional como archivistas, podemos caracterizar los errores y omisiones que se identifiquen para sugerir si es necesario ampliar el conjunto de entrenamiento para subsanarlos en posteriores versiones de la IA.

Un caso parecido para la gestión de documentos se desarrolló en Australia. En noviembre y diciembre de 2017 se documentó una prueba piloto en los Archivos Estatales de Nueva Gales del Sur (NSWSAR) realizada por el equipo de Archivos Digitales. El objetivo principal era probar algoritmos de aprendizaje automático en un corpus de documentos que previamente habían sido valorados manualmente por una autoridad para su baja documental. La pregunta era: “¿Con qué nivel de precisión se podría realizar una clasificación de múltiples valores para cotejar automáticamente el corpus con las mismas

clases de baja?” (Rolan et al., 2018, p. 191). Ejercicios como este, al trasladar criterios de valoración documental a valores cuantitativos identificables con el procesamiento de información, permite reconocer los alcances y límites de la IA.

El proyecto decidió no utilizar ninguna solución basada en la nube, ya que no podría cumplir del todo con la legislación aplicable o requeriría una asesoría legal profunda, pero el plazo del proyecto impedía una evaluación de riesgos más detallada. Por esta razón se buscó un software disponible que pudiera ejecutarse localmente, ya que el proyecto piloto interno incluía un corpus de documentos cerrados que habían sido transferidos como Archivos Estatales (Rolan et al., 2018, p. 191).

A continuación, se incluye información sobre el preprocesamiento de datos de este caso de estudio porque nos permitirá ilustrar labores en las que mediante nuestra profesión podemos sumarnos durante el desarrollo de modelos de IA.

El corpus completo comprendía 30 GB de datos, en 7,561 carpetas, que contenían 42,653 expedientes; no se había aplicado ninguna norma de eliminación a los expedientes. En un esfuerzo conjunto con el departamento, el corpus fue valorado manualmente (a nivel de carpetas) con respecto a los Registros Administrativos de la Autoridad General de Retención y Eliminación (GA28), lo que dio como resultado un total de 12,369 expedientes requeridos como Archivos del Estado, que fueron preprocesados con el siguiente flujo (Rolan et al., 2018, pp. 192-193):

- **Extracción de texto:** Sólo se seleccionaron para su uso 8,784 de los 12,369 expedientes clasificados como necesarios como archivos estatales. Esto se debió a que sus tipos de documento permitían una extracción de texto sencilla. Se desarrolló un programa en Python que utilizaba varias bibliotecas para extraer texto de los tipos de archivo PDF, DOCX y DOC. El texto extraído de los documentos se colocó en un único archivo de valores separados por comas (CSV).

- **Limpieza de datos:** Se aplicaron los siguientes filtros: eliminación del formato de los documentos; eliminación de las palabras vacías, eliminación de los documentos que no son necesarios y conversión de todas las letras a minúsculas.
- **Extracción de características:** Se trata de un modelo sencillo que no tiene en cuenta la ubicación de las palabras en los documentos, sino que se centra en la presencia y la frecuencia de las palabras individuales. Este enfoque puede representar cualquier documento como un conjunto de palabras únicas.
- **Clasificación:** Para la clasificación del texto se utilizaron algoritmos estadísticos y matemáticos

Cabe destacar la importancia en que el entrenamiento de los datos resulta fundamental para una correcta precisión del algoritmo (Rolan et al., 2018, p. 193).

El flujo anterior de trabajo no tiene que ser necesariamente conocido a profundidad pero si debe existir cierta vigilancia en que se lleven del mejor modo posible, para generar resultados de mejor calidad. Durante el desarrollo de este flujo de trabajo podemos contribuir con la selección de documentos inicial; en el paso de clasificación se vuelve a actuar para analizar los resultados.

Las pruebas sugieren que, si bien la IA promete eficiencia en el apoyo a la valoración digital, la documentación y la eliminación, la tasa de éxito fue del 84%, lo que en el estudio resulta inferior al compararse con las posibilidades humanas (Rolan et al., 2018, p. 195).

Esto que se lee como un indicio de que ciertos procesos como la de valoración documental pueden aprovechar esta herramienta a modo de primer contacto, produciendo una visión general de los documentos, pero en la decisión final debería priorizarse nuestra opinión como archivistas, siendo esta tecnología solamente un apoyo más en el procesamiento de grandes volúmenes de documentos. En este estudio, por mencionar solo algunas actividades es posible involucrarnos en la documentación del sistema desde un inicio y durante su desarrollo, además de retroalimentar los resultados de la IA. Al final de

este capítulo se ofrecen una serie de actividades iniciales como punto de partida para la integración del archivista en este proceso.

Utilizar estas tecnologías permite optimizar el tiempo, pero es importante garantizar que la responsabilidad de nuestros sistemas y procesos no se vea comprometida (Rolan et al., 2018, p. 195).

Otra de las pruebas realizadas con procesamiento del lenguaje natural fue en el Consejo Europeo en donde se tenía la problemática de describir 1,000,000 de documentos a nivel de unidad simple (Büttner., 2019, pp.1-2). Este proceso podría ser parcialmente automatizado, dado que los documentos eran nacidos en digital o digitalizados y se tomaban valores predefinidos a nivel de unidad documental compuesta o serie, pero esto no generaba metadatos útiles de descripción para la recuperación de documentos de unidad simple, por lo que se decidió utilizar otro enfoque basado de clasificación automatizada para capturar la idea del texto y localiza entidades como personas, lugares u organizaciones (Büttner., 2019, pp.1-2),

Se utilizó un software para realizar una prueba de concepto en el uso de 10,000 documentos los cuales fueron tomados de una amplia procedencia de cuerpos intergubernamentales, que a su vez trataban diferentes asuntos como derechos humanos, justicia criminal, herencia cultural entre otros. Una limitante que se halló mediante el uso del software, fue la identificación de documentos donde el nombre de un país estaba precedido por un nombre de algún representante, generando una errónea identificación ya que el software lo identificaba como si el documento tratara acerca de un país. Por esta razón, se creó una regla general para ignorar esos casos (Büttner, 2019, pp.19-21).

Se realizó un estudio con 13 usuarios de 7 sectores diferentes, sobre el cual se midieron las siguientes 3 métricas de rendimiento: 1) Búsqueda utilizando los metadatos creados por el algoritmo, que resultó en una mejora del 21% en tanto a identificar resultados relevantes. 2) La precisión de la búsqueda, que mejoró el número de documentos relevantes en la primera página usando refinadores, en un 18%. 3) Los metadatos se midieron en

relación con la información adecuada y correcta con respecto al contenido de un documento; para el 85,5% se calificó como media a alta. Por otro lado, se decidió establecer los metadatos creados por los profesionales de la información como el estándar a seguir, dado que el proceso para su realización era más amplio en comparación al realizado por el algoritmo (Büttner, 2019, pp. 21-25). La IA pudo demostrar en este caso una mejora en el rendimiento, pero es prudente insistir que en este caso los metadatos creados pueden brindar una base, que posteriormente debería ser enriquecida en el marco de una descripción archivística.

Más allá de los documentos administrativos, fiscales, legales o contables que normalmente gestionamos, hay ocasiones en que los documentos son de otra clase y en ellos está contenida la opinión pública, como lo pueden ser las encuestas. Una manera de procesar esta gran cantidad de información es aprovechando la IA pero además de ello, poder enriquecer el posible análisis de estos documentos con la experiencia archivística puede ampliar la perspectiva de análisis. Tal es la propuesta de Marie-Anne Chabin para el caso de la opinión pública del Gran Debate Nacional en Francia, llevada a cabo en los primeros meses del año 2019.

El "gran debate nacional" recopiló datos a través de dos iniciativas paralelas como sigue: la primera desde diciembre del 2018 fue en todo el país donde mediante reuniones con la población se podía expresar espontáneamente sus preocupaciones a los cargos electos y al gobierno. La segunda iniciativa consistió en la inclusión de las contribuciones de la población en general en una plataforma nacional, lanzada oficialmente a mediados de enero de 2019 y estructurada en torno a cuatro temas predefinidos, a saber, la transición ecológica, la economía y el gasto público, la democracia y la ciudadanía y los servicios públicos.

Los cuestionarios de los cuatro temas sumaban 84 preguntas, que incluían no sólo preguntas cerradas (sí/no o una lista de respuestas), sino también preguntas abiertas y otras que fomentan el desarrollo de una idea. El material documental fue el siguiente: 20000

cuadernos, 27000 cartas y correos, 1.9 millones de contribuciones en la plataforma y 20 reportes de conferencias nacionales y regionales. (Chabin, 2020, pp. 242-243.) En este caso el volumen de información recogida era extenso, pero aquella recolectada en la plataforma -al ser su formato de origen digital-, podía ser analizada con mayor facilidad mediante el uso de la IA.

Todo el material reunido se procesó por dos grupos diferentes. La primera parte fue analizada por una empresa privada, que analizó las opiniones de la plataforma y elaboró cuatro síntesis, una por cada uno de los cuatro temas principales de la plataforma: ecología, fiscalidad, democracia y Estado y servicios públicos. Los conceptos citados por los encuestados fueron identificados, analizados, ordenados y clasificados en diferentes categorías y subcategorías. A continuación, se reindexó la base de datos de textos y se asoció el resultante de cada pregunta abierta a una o varias categorías. Por último, una nueva operación de recuento permitió jerarquizar los temas y producir porcentajes. (Chabin, 2020, p. 244.)

Mediante el uso de algoritmos de AA es posible agrupar conceptos, en forma de categorías, pero es conveniente señalar que un análisis semejante también puede ser enriquecido con la procedencia de los datos y sumar una capa de análisis respecto a la ubicación geográfica de procedencia.

Para el caso del material externo a la plataforma, primero se dio prioridad a las "propuestas" concretas de los ciudadanos para resolver los problemas. Después, el tratamiento se extendió a las "preocupaciones" de la población. Este marco de referencia se estructuró entonces en 8 temas (además de los cuatro temas definidos por el gobierno, se añadieron otros cuatro temas surgidos de los análisis: economía y empleo; poder adquisitivo; educación y formación; y salud, solidaridad e integración). Para realizar el contraste entre el marco temático de referencia y las opiniones recolectadas, los criterios organizativos de éstas fueron los temas abordados y el análisis de las preocupaciones; la tipología de las

contribuciones (cuadernos, cartas y actas) y los criterios geográficos también se analizaron cuantitativamente (Chabin, 2020, pp. 244-245.)

Chabin nos indica dos modos de procesamiento de datos con base en la IA siendo el primero semántico de tipo supervisado: el tratamiento parte de categorías predefinidas y busca en el corpus los elementos que pueden adscribirse a ellas. La principal crítica a este enfoque es que, al presuponer los temas tratados en el corpus, podemos estar privándonos de descubrimientos interesantes. En lo que respecta al gran debate, las sugerencias originales y/o relevantes pueden pasar desapercibidas porque no se recogerán a través de temas predefinidos.

El segundo, de tipo no-supervisado, consistió en utilizar fórmulas matemáticas para considerar la combinación de proximidad y similitud, que es parte fundamental de investigaciones en ciencias cognitivas. Este enfoque permite un análisis más exhaustivo del contenido real del corpus, a la vez que ofrece la posibilidad de identificar formulaciones o afirmaciones que van más allá de lo esperado o asumido por los autores del corpus, ideas que podrían permanecer ocultas (Chabin, 2020, p. 246.). Las decisiones tomadas en este caso son especialmente interesantes para mostrar la conveniencia de combinar ambos tipos de aprendizaje automático.

Ambos enfoques no están enfrentados, sino que pueden ser complementarios: el semántico se podría haber implementado para realizar un análisis con base en las preguntas de la investigación y el no supervisado para identificar nuevas ideas no identificadas con el primero. Por otra parte cabe destacar, según Chabin, que la IA se basa en el texto pero no abarca, el contexto en el que fueron producidos los documentos: lugar, fecha, diseño, estilo o relación con un grupo de contribución (Chabin, 2020, p. 246.). Este modelo de IA tiene como limitante que su funcionamiento regular se base en el análisis de diferentes conjuntos de texto, si se enriquece el conjunto con metadatos que proporcionen contexto, podría resultar en un análisis más adecuado.

La diplomática posibilitaría tener en cuenta el entorno de producción y la forma en que se expresaron los contribuyentes, con la intención de demostrar el grado de fiabilidad de este material. Para responder preguntas como "¿de qué es representativa exactamente esta consulta gubernamental? o ¿en qué sentido no es representativa?", sería pertinente no únicamente analizar de lo que se ha expresado a través de palabras y frases parciales, sino también más allá de las palabras incorporando del contexto de producción. (Chabin, 2020, pp. 247-248).

En el caso de la archivística, el uso de la diplomática puede caracterizar el documento digital que es recibido por la IA, teniendo en cuenta su contexto de producción. Es decir, desde el punto de vista del procesamiento de datos, puede brindar más datos de contexto. Lo propuesto por Chabin es un acercamiento inicial, este trabajo abre un horizonte en términos de qué otros elementos es posible incorporar para el procesamiento a gran escala. Haría falta en el futuro la formulación de una metodología que pudiera enriquecer este cruce de datos desde la archivística y hacia la IA

Además del contenido particular de cada tipo de documento; para el caso de grandes volúmenes documentales -como éste-, la diplomática puede aportar criterios de estudio y contexto al momento de la creación de conjuntos de datos de entrenamiento.

Dentro de los resultados, las contribuciones al debate nacional se ubicaron en 3 niveles. El primer nivel de opinión individual en forma de párrafo, carta o respuesta en línea. El segundo corresponde al conjunto de contribuciones que pueden ser, las libretas de quejas, minutas o un conjunto de respuestas de la plataforma. El tercero podría identificarse a partir de criterios temáticos y/o geográficos externos a los contribuyentes (por ejemplo, los cuadernos de una región, la serie de respuestas a uno u otro de los cuatro temas oficiales).

La integración de elementos diplomáticos para enriquecer el tratamiento algorítmico, con la identificación de los diferentes niveles de objetos documentales, habría permitido, por ejemplo, extraer del material inicial otros elementos de conocimiento sobre la población de contribuyentes y el comportamiento de sus diferentes perfiles (Chabin, 2020, p. 250).

Además, se observa que las opiniones en plataforma tenían un campo denominado título, pero este no se consideró en el uso de la IA. Finalmente, varias preguntas de la encuesta no fueron respondidas. Lo anterior implica un vacío en la respuesta que también tendría un significado para pensarse, hecho faltante en este estudio (Chabin, 2020, p. 249). Estas observaciones proporcionan ejemplos concretos de formas en que el estudio documental podría enriquecer el análisis computacional.

Como reflexión final a partir de esta aproximación a la IA, se insiste en la importancia de potenciar las interacciones entre las tecnologías de tratamiento de la información y los conceptos probados de análisis de documentos. Es fundamental desarrollar proyectos junto con las empresas tecnológicas y los proveedores de inteligencia artificial para identificar el conocimiento social y el de los documentos.

La primera fase consistió en el estudio cualitativo de 8 profesionales en la gestión de la información que incluyó entrevistas, pruebas de valoración de correos electrónicos, los correos electrónicos seleccionados y la posterior interpretación de los resultados, a partir de los cuales fue posible establecer un patrón en común. Parte del entrenamiento del algoritmo consistió en estudiar el lenguaje utilizado en los correos, donde se conformaron cuatro dimensiones: social, pragmática, temporal y semántica. Se llegó a un consenso donde el valor empresarial pudo ser comprendido en el contexto de la rendición de cuentas, como los documentos sobre la captación de clientes. (Vellino y Alberts., 2016, pp.298-299). Esta es una clara muestra de que en cada archivo la forma en que se realiza la valoración cambia según el contexto, por lo que la implementación de una IA de carácter universal podría no ser efectiva sin una investigación en el contexto determinado.

Establecer el valor empresarial es algo que debe estar relacionado con los integrantes de la organización y que tiene como factor la cultura. Por lo anterior, es conveniente tomar en cuenta metodologías para el entrenamiento de los clasificadores para cuestiones relacionadas a la gestión documental. Una particularidad acerca del uso de AA fue que los errores cometidos por el algoritmo eran consistentes, a diferencia de las variaciones que

dependen de las personas que realizan la valoración y el periodo en el cual lo efectúan (Vellino y Alberts., 2016, pp. 306-309). Este es un ejemplo de cómo los archivistas podemos participar en el diseño de la IA, así como en su posterior evaluación; una forma de promover esto es la discusión de los errores consistentes por equipos propios donde podríamos sugerir un conjunto adicional de documentos de entrenamiento que permitan su posterior corrección, colaborando así con el equipo de desarrolladores.

A partir de este caso de estudio, se reconoce como necesario incorporar el conocimiento sobre el contexto organizacional y los requerimientos de los usuarios finales al entrenar al algoritmo; sabiendo que la IA siempre necesitará del contexto para operar del mejor modo, como archivistas podemos hacer aportaciones valiosas para incorporar esta información contextual.

Para establecer el control sobre los conjuntos de datos, es necesario tener un rol activo en el entrenamiento de un algoritmo de AA. Hutchinson (2018) presenta un avance preliminar de un caso de estudio del uso de aprendizaje supervisado y su conexión con la asistencia para la identificación de documentos de recursos humanos mediante el modelado de tópicos, donde utilizó Weka¹⁴, un software de minería de datos, para analizar el corpus de texto el cual está integrado por los documentos del Vicepresidente Asociado de Tecnologías de la Información y la Comunicación de la Universidad de Saskatchewan. (Hutchinson, 2018, p. 2696-2697). El corpus según (Hutchinson, 2018) estaba clasificado en:

- **Documentos personales de recursos humanos (RH-personal):** Información de recursos humanos relativos a uno o más individuos identificados.
- **Documentos generales de recursos humanos (RH-general):** Documentos relacionados con los recursos humanos pero que no identifican a individuos, como

¹⁴ Weka (Waikato Environment for Knowledge Analysis) es una plataforma de software destinada al aprendizaje automático y la minería de datos. Está escrito en lenguaje Java y fue desarrollado en la Universidad de Waikato en Nueva Zelanda

descripciones de puestos de trabajo, las políticas y la correspondencia y actas relacionadas con su desarrollo, etc.

- **Documentos no relacionados con los recursos humanos (no RH):** todos los demás registros.

El enfoque del estudio fue realizar varias combinaciones dentro del conjunto de documentos como forma de comprender qué clase de selección daba como resultado una mejor clasificación. En la práctica, durante la selección de documentos para el desarrollo de los conjuntos de entrenamiento se escogieron de forma parcial o total los documentos del Vicepresidente Asociado de Tecnologías de la Información y la Comunicación, se removieron aquellos tipos documentales que no contuvieran información personal relevante, se eligieron aquellos documentos con mayor extensión y, finalmente, se configuró el software Weka de varias formas para analizar los resultados (Hutchinson, 2018, pp.2697-2698.)

Como resultado del estudio se observan las posibles ventajas para usar aprendizaje supervisado para identificar documentos con datos personales o establecer un posible riesgo. Según (Hutchinson, 2018) proponen también direcciones futuras de investigación:

- Realizar pruebas con otra clase de clasificadores.
- Establecer pruebas más específicas con tipos documentales específicos.
- Analizar los falsos negativos y falsos positivos y buscar sus causas.
- Desarrollar conjuntos de entrenamiento para identificar documentos de acceso restringido o no restringido.

Este caso de estudio muestra cómo la IA en su uso para la agrupación puede mostrar diferentes resultados según la información suministrada. No es trivial la selección de distintas clases de documentos para entrenar a una IA, ya que el conjunto de documentos empleados, determina la utilidad de su funcionalidad. La importancia de robustecer la creación de conjuntos de entrenamiento es uno de los aspectos que se señalan con mayor coincidencia en los trabajos recuperados dentro de esta revisión, con énfasis en la

documentación que se puede generar sobre ello (Rolan et al., 2018; Andresen, 2019; Gebru, 2020).

A partir de la revisión de literatura, incluyendo los casos seleccionados, es posible identificar actividades que posteriormente se pueden agrupar en funciones que, como archivistas, podemos cumplir en relación con la IA. En la Tabla 6 se muestran las funciones que a partir de la lectura de los casos de estudios mencionados, en los que podemos tener una participación activa desde la archivística.

Funciones	Actividades
Documentar	<ul style="list-style-type: none"> • Documentación del proceso de implementación de la IA (Andresen, 2019; Gebru,2020; Rolan et al., 2018). • Generar aquella documentación necesaria que sirva para explicar el algoritmo y los datos de su entrenamiento (Andresen, 2019; Bunn, 2020; Gebru, 2020; Hutchinson, 2018).
Participación en el desarrollo de modelos	<ul style="list-style-type: none"> • Involucrarse en el desarrollo desde el proceso de entrenamiento del algoritmo (Jailliant y Caputo,2022; Vellino & Alberts, 2016). • Sugerir la consideración de determinados metadatos que enriquezcan el análisis por parte del algoritmo (Chabin, 2020; Makhoul-Shabou et al., 2020). • Establecer los tipos documentales adecuados para el entrenamiento (Makhoul-Shabou et al., 2020). • Contribuir en la selección de palabras clave para la minería de texto (Hutchinson,2020). • Identificar omisiones o falsos positivos del algoritmo (Büttner, 2019; Hutchinson, 2018, 2020; Gebru, 2020)
Ética profesional	<ul style="list-style-type: none"> • Vigilancia del seguimiento de principios éticos y transparencia (Gebru, 2020); Andresen, 2019; Colavizza et al., 2022; Jailliant y Caputo, 2022). • Implicar la gestión de riesgos a la IA (Andresen, 2019). • Análisis del uso del software o algoritmo con base en los requerimientos legales aplicables (Andresen, 2019; Hutchinson, 2018; Jailliant y Caputo, 2022; Rolan et al., 2018). • Aprender de forma superficial la IA mediante guías para obtener una mayor comprensión de su funcionamiento (Jailliant y Caputo, 2022).
Pertinencia y mejora	<ul style="list-style-type: none"> • Análisis crítico del rendimiento, características y mejoras de la IA en su implementación (Büttner,2019; Hutchinson, 2018; Makhoul-Shabou et al., 2020). • Sugerir mejoras al software en su usabilidad (Hutchinson, 2020). • Generar requerimientos específicos para el desarrollo del software tomando en cuenta las necesidades del archivista como por ejemplo: el cuadro de clasificación (Hutchinson, 2020).
Análisis crítico situado	<ul style="list-style-type: none"> • Valorar el uso de determinados algoritmos para analizar un conjunto de documentos en específico (Chabin, 2020). • Formar parte de la discusión de la explicabilidad de la IA (Bunn, 2020; Jailliant y Caputo, 2022). • Identificar elementos diplomáticos o archivísticos que ayuden al análisis de las limitaciones del uso de un algoritmo específico (Chabin, 2020; Gebru, 2020).

Tabla 7. Funciones y acciones del archivista en relación con la IA en general Fuente: Elaboración propia

En el contexto de esta investigación cada término que agrupa las actividades que significan lo siguiente:

- **Documentación:** Utilizar los instrumentos necesarios para la evaluación y seguimiento del desarrollo de la IA, así como resguardar y dar acceso a esta información.
- **Participación en el desarrollo de modelos:** Involucrarse en acciones que influyan directamente en el desarrollo de la IA.
- **Ética profesional:** Supervisar el desarrollo de la IA mediante la vigilancia de principios éticos y de transparencia, la gestión de riesgos y las implicaciones regulatorias aplicables
- **Pertinencia y mejora:** A partir del uso específico de la IA en el archivo, especificar nuevas funciones para el diseño de la IA, o la mejora de las ya existentes.
- **Análisis crítico situado:** Analizar cómo la IA transforma los procesos archivísticos, para realizar un desarrollo informado sobre las particularidades de cada contexto de aplicación.

Las acciones mencionadas nos colocan como un actor relevante en el diseño e implementación de esta tecnología; como se ha visto, esto representa un gran desafío, ya que la IA abarca diferentes áreas de actuación que implican realizar un análisis robusto de diferentes componentes, ya sean tecnológicos, sociales, archivísticos o legales, según cada caso concreto.

En este capítulo se abordó un panorama sobre la IA en el archivo, donde principalmente los casos presentados fueron sobre: identificación, descripción y valoración de documentos. La organización es una de nuestras principales funciones, y con este se tiene un acercamiento a cómo la IA complementa nuestro tratamiento de información en grandes volúmenes. Con nuestra experiencia y conocimiento podemos colaborar en el

diseño, implementación y mejora de una IA más robusta que nos fortalezca profesionalmente.

La protección de datos personales frente a la inteligencia artificial y su fiabilidad desde la archivística

Las normativas de transparencia y privacidad en México

Derivado de la expansión en las últimas décadas de las tecnologías de información y comunicaciones, se han conformado dinámicas de tratamiento, producción y gestión de la información en entornos digitales, cada vez con mayor automatización. Esto brinda a gobiernos y organizaciones un medio computarizado para contactar y recabar información para distintos trámites; en ocasiones, incluso cuando la gestión se realiza de manera analógica por medio del papel, posteriormente es digitalizada para su preservación, conservación y consulta.

Tal como se describió en los casos anteriores, los archivos digitales son un lugar natural para la implementación de diferentes aplicaciones de IA, como el procesamiento del lenguaje natural, el procesamiento de la imagen y del sonido. Una de las posibles esferas de aplicación es la protección de datos personales. Cualquier organización que posea información, está expuesta a riesgos potenciales como filtraciones, robos y ataques informáticos. Estos peligros afectan tanto a entidades gubernamentales como a empresas privadas. Hay un interés en el acceso indebido a documentos que contienen datos personales, ya que esto puede conducir a la suplantación de identidad, la extorsión, la discriminación, el chantaje o la amenaza, por parte de actores mal intencionados.

En México los antecedentes legislativos sobre la protección de datos personales se encuentran en la Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental del 11 de junio de 2002¹⁵, que en términos generales sienta las bases en

¹⁵ Cámara de diputados del H. Congreso de la Unión. Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental Disponible en: <https://www.uv.mx/contraloria/files/2013/02/8.-Ley-Federal-de-Transparencia-y-Acceso-a-la-Informacion-Publica-Gubernamental.pdf>

relación de la regulación al acceso a la información gubernamental y describe en menor medida la protección de datos personales.

Años más tarde, en 2010, se aprobó la Ley Federal de Protección de Datos Personales en Posesión de los Particulares¹⁶, que se centra específicamente en los datos personales enfocándose en varios aspectos como el tratamiento que los particulares (personas físicas y morales) dan a los datos personales de las personas, para proteger su privacidad. En el año 2017 se aprobó la Ley General de Protección de Datos Personales en Posesión de Sujetos Obligados¹⁷ (LGPDPPSO), que es la última versión publicada hasta el momento. Con ella, aparece una mejora en los siguientes aspectos:

- Establece mayores obligaciones y responsabilidades para los sujetos obligados, como designar un responsable de protección de datos, mantener registro de actividades de tratamiento y realizar evaluaciones de impacto en la protección de datos.
- Otorga más derechos a los titulares de los datos personales, como el de acceso, rectificación, cancelación y oposición (ARCO) o el derecho a la portabilidad de datos, por mencionar algunos
- Prevé un régimen más robusto de sanciones por incumplimiento de las disposiciones de la ley, que pueden ser multas, amonestaciones públicas, bloqueo de datos personales, etc.

Con el paso del tiempo se ha tenido más conciencia acerca de la importancia de la privacidad de forma legalmente vinculante. Además, se ha puesto una especial área de atención, al control de las nuevas tecnologías de la información tanto en el ámbito público como en el privado, para prevenir el mal tratamiento de los datos personales.

¹⁶ Cámara de diputados del H. Congreso de la Unión. Ley Federal de Protección de Datos Personales en Posesión de los Particulares. Disponible en: <https://www.diputados.gob.mx/LeyesBiblio/pdf/LFPDPPP.pdf>

¹⁷ Cámara de diputados del H. Congreso de la Unión. Ley General de Protección de Datos Personales en Posesión de Sujetos Obligados Disponible en: https://www.diputados.gob.mx/LeyesBiblio/ref/dof/CPEUM_ref_208_11jun13.pdf

En la actualidad hay varias leyes que conforman un marco legislativo que promueve la transparencia. En relación a los archivos, está la primera Ley de Archivos en México, como parte del ejercicio de la transparencia en el ámbito público, conocida como Ley Federal de Archivos, publicada en 2012. El principal objetivo de esta ley es controlar la forma en que la administración pública federal de México organiza y conserva sus archivos. Además, esta Ley enuncia como objetivos en el artículo 6º los siguientes: (Congreso de la Unión, 2012, 5)

- Establecer disposiciones para garantizar el derecho a la verdad y el acceso a la información contenida en los archivos públicos.
- Establecer los procedimientos y responsabilidades para la creación, uso, preservación y valoración de los documentos públicos.
- Establecer las bases para organizar, administrar, preservar y proteger el patrimonio documental de la nación.
- Promover la cultura archivística y el uso de tecnologías de información en la gestión de documentos.

Somos esenciales para promover y fomentar la transparencia, ya que unos archivos debidamente organizados y gestionados permiten a las organizaciones y a las personas consultar la información; sin esta labor, la transparencia se vería comprometida y la capacidad de las personas para supervisar y exigir responsabilidades a los gobiernos sería limitada. Tal como señalan Cerrillo Martínez y Casadesús de Mingo (2018, p. 14), los registros y la información que cada administración pública crea y mantiene en el ejercicio de sus funciones son objeto de transparencia.

En 2015 se aprueba la Ley General de Transparencia y Acceso a la Información Pública¹⁸ (LGTAIP) que establece el derecho de cualquier persona a solicitar y acceder a la información pública generada por las instituciones gubernamentales, y obliga a todas las

¹⁸ Cámara de diputados del H. Congreso de la Unión. Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental Disponible en: https://www.diputados.gob.mx/LeyesBiblio/pdf/LGTAIP_200521.pdf

dependencias del gobierno mexicano a rendir cuentas sobre su información y ponerla a disposición de los ciudadanos a través de portales de transparencia.

En 2018, se aprobó la Ley General de Archivos, que a grandes rasgos regula la organización y conservación de archivos en posesión de cualquier autoridad, entidad, órgano y organismo que reciba y ejerza recursos públicos a nivel federal, estatal y municipal. Entre los diferentes objetivos hallados en el artículo 2, destaco la siguiente fracción IV:

Promover el uso y aprovechamiento de tecnologías de la información para mejorar la administración de los archivos por los sujetos obligados. (Congreso de la Unión, 2018, 2)

Con ello, entra dentro nuestra competencia el aprovechamiento de la tecnología disponible para realizar mejoras en la gestión, la disponibilidad, la consulta, la difusión y la organización de los documentos. A su vez también los archivistas tenemos las siguientes funciones, según lo estipula el artículo 42:

Los sujetos obligados establecerán en su programa anual los procedimientos para la generación, administración, uso, control y migración de formatos electrónicos, así como planes de preservación y conservación de largo plazo que contemplen la migración, la emulación o cualquier otro método de preservación y conservación de los documentos de archivo electrónicos, apoyándose en las disposiciones emanadas del Consejo Nacional. (Congreso de la Unión, 2018, 16)

De lo anterior se puede interpretar que, para la gestión de documentos electrónicos, debe haber políticas y estrategias para asegurar que los archivos electrónicos permanezcan íntegros, accesibles y utilizables a largo plazo. Esto es particularmente relevante en el caso de los archivos digitales, ya que se deben mantener vigentes incorporando las nuevas tecnologías de la información que puedan contribuir a estos procesos

Cabe destacar que de igual forma en México en el artículo 6.º constitucional se establece que el Estado está obligado a garantizar el derecho a la información . En lo que respecta a los archivos hay una mención especial en este mismo artículo, en la fracción reformada del año 2014 en la fracción V del inciso A:

Los sujetos obligados deberán preservar sus documentos en archivos administrativos actualizados y publicarán, a través de los medios electrónicos disponibles, la información completa y actualizada sobre el ejercicio de los recursos públicos y los indicadores que permitan rendir cuenta del cumplimiento de sus objetivos y de los resultados obtenidos (Congreso de la Unión, 2014, 13) .

El artículo anterior determina medidas para garantizar la transparencia y la rendición de cuentas en la gestión pública, teniendo como eje los archivos, para promover la confianza en las instituciones. Los medios electrónicos para la publicación de información permiten una amplia difusión y fácil acceso, habilitando a la sociedad evaluar y participar, siendo los archivos parte esencial. En este artículo se demuestra ese reconocimiento.

[Los datos personales, las versiones públicas y la anonimización](#)

El artículo 3º, fracción IX, de la LGPDPSO, define a los datos personales como:

Cualquier información concerniente a una persona física identificada o identificable. Se considera que una persona es identificable cuando su identidad pueda determinarse directa o indirectamente a través de cualquier información (Congreso de la Unión, 2017, 3)

En ese mismo artículo en la fracción X se definen a los datos personales como:

Aquellos que se refieran a la esfera más íntima de su titular, o cuya utilización indebida pueda dar origen a discriminación o conlleve un riesgo grave para éste. De manera enunciativa más no limitativa, se consideran sensibles los datos personales que puedan revelar aspectos como origen racial o étnico, estado de salud presente o futuro, información genética, creencias religiosas, filosóficas y morales, opiniones políticas y preferencia sexual.(Congreso de la Unión, 2017, 13)

Estos datos deben ser protegidos siendo que puede violarse la privacidad de las personas, al ser información personal que está dentro de la esfera privada. Para ello, aquella información que resulta tener datos personales sensibles, puede ser clasificada como reservada o confidencial; en la Ley General de Transparencia y Acceso a la Información Pública se define como clasificación en el artículo 100º:

La clasificación es el proceso mediante el cual el sujeto obligado determina que la información en su poder actualiza alguno de los supuestos de reserva o confidencialidad, de conformidad con lo dispuesto en el presente Título (Congreso de la Unión, 2015, 39).

Como resultado de la solicitud de información se brindan documentos, los cuales pueden tener párrafos que contengan datos confidenciales o reservados; en ese caso, estos deben ser testados tal como lo indica el artículo 111º de dicha Ley:

Cuando un Documento contenga partes o secciones reservadas o confidenciales, los sujetos obligados, para efectos de atender una solicitud de información, deberán elaborar una Versión Pública en la que se testen las partes o secciones clasificadas, indicando su contenido de manera genérica y fundando y motivando su clasificación (Congreso de la Unión, 2015, 41).

En el artículo 113º de dicha Ley se indican como contenidos reservados aquellos que pueden comprometer la seguridad nacional, las relaciones internacionales, el sistema financiero, o bien obstruir actividades de persecución de delitos, así como la información de procesos deliberativos o judiciales en curso, por mencionar algunos (Congreso de la Unión, 2015, 42). En cuanto a la información confidencial en el artículo 116º se dice lo siguiente: “Se considera información confidencial la que contiene datos personales concernientes a una persona identificada o identificable.” (Congreso de la Unión, 2018, 42). En el mismo artículo se especifican qué clase de información podría ser sujeta de esta, se dice:

Se considera como información confidencial: los secretos bancario, fiduciario, industrial, comercial, fiscal, bursátil y postal, cuya titularidad corresponda a particulares, sujetos de derecho internacional o a sujetos obligados cuando no involucren el ejercicio de recursos públicos.

Asimismo, será información confidencial aquella que presenten los particulares a los sujetos obligados, siempre que tengan el derecho a ello, de conformidad con lo dispuesto por las leyes o los tratados internacionales (Congreso de la Unión, 2015, 42) .

Según el artículo 129º de esta misma Ley:

Los sujetos obligados deberán otorgar acceso a los Documentos que se encuentren en sus archivos o que estén obligados a documentar de acuerdo con sus facultades, competencias o funciones en el formato en que el solicitante manifieste, de entre aquellos formatos existentes, conforme a las características físicas de la información o del lugar donde se encuentre así lo permita (Congreso de la Unión, 2015, 45).

En algunas ocasiones no será posible que el solicitante puede tener acceso completo al contenido del documento, ya que este puede contener partes confidenciales o reservadas, de modo que, para brindar el acceso a la información según lo dicta la misma Ley en el artículo 111º, se tendrá que generar una versión pública donde se testen las partes necesarias, es decir, donde se eliminan u omiten las secciones clasificadas (Congreso de la Unión, 2015, 45)

Las versiones públicas son un medio para generar un equilibrio entre la transparencia y la protección de datos personales. Los Lineamientos generales en materia de clasificación y desclasificación de la información, así como para la elaboración de versiones públicas¹⁹, definen testar como: "La omisión o supresión de la información clasificada como reservada o confidencial, empleando sistemas o medios que impidan la recuperación o visualización de ésta" (Consejo Nacional del Sistema Nacional de Transparencia, Acceso a la Información Pública y Protección de Datos Personales, 2016). Tales lineamientos tienen un tratamiento diferente para el caso del formato del documento ya sea impreso o digital. En el caso de la versión impresa se establece lo siguiente:

Quincuagésimo noveno. En caso de que el documento únicamente se posea en versión impresa, deberá fotocopiar y sobre éste deberán testarse las palabras, párrafos o renglones que sean clasificados, debiendo anotar al lado del texto omitido, una referencia numérica tal y como se puede observar en el modelo para testar documentos impresos contenido en el Anexo 1 de los Lineamientos, "Modelo para testar documentos impresos" (Consejo Nacional del Sistema Nacional de Transparencia, Acceso a la Información Pública y Protección de Datos Personales, 2016).

En la ilustración 1 se muestra un ejemplo de un documento impreso testado.

¹⁹ ACUERDO del Consejo Nacional del Sistema Nacional de Transparencia, Acceso a la Información Pública y Protección de Datos Personales, por el que se aprueban los Lineamientos generales en materia de clasificación y desclasificación de la información, así como para la elaboración de versiones públicas. Publicado el 15 de abril del 2016. Disponible en: https://www.dof.gob.mx/nota_detalle.php?codigo=5433280&fecha=15/04/2016#gsc.tab=0

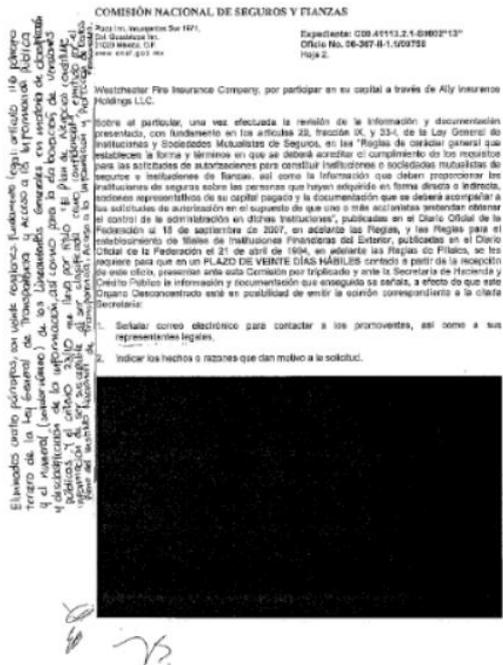


Ilustración 1 Modelo para testar documentos impresos [Ilustración], por Consejo Nacional del Sistema Nacional de Transparencia, Acceso a la Información Pública y Protección de Datos Personales, 2016, Lineamientos generales en materia de clasificación y desclasificación de la información, así como para la elaboración de versiones públicas, (https://www.dof.gob.mx/nota_detalle.php?codigo=5433280&fecha=15/04/2016#qsc.tab=0)

Para el caso de la versión digital se establece lo siguiente:

Sexagésimo primero. En la parte del documento donde se hubiese ubicado originalmente el texto eliminado, deberá insertarse un cuadro de texto en color distinto al utilizado en el resto del documento con la palabra "Eliminado", el tipo de dato o información cancelado y señalarse si la omisión es una palabra(s), renglón(es) o párrafo(s) (Consejo Nacional del Sistema Nacional de Transparencia, Acceso a la Información Pública y Protección de Datos Personales, 2016).

En la ilustración 2 se muestra un ejemplo para el testado de documentos en formato electrónico



Fecha de Clasificación: 25 de junio de 2005.
 Unidad Administrativa: Dirección General de
 Clasificación de Información y Datos Personales.
 Reservada. Plazo única.
 Período de reserva: Dos años.
 Fundamento Legal: Artículo 14 fracción III
 L.FTAMP.
 Ampliación del período de reserva:
 Confidencial: X X X
 Fundamento Legal:
 Rúbrica del titular de la Unidad Administrativa.
 Fecha de desclasificación:
 Rúbrica y cargo del servidor público.

INSTITUTO FEDERAL DE ACCESO A LA INFORMACIÓN PÚBLICA
 DIRECCIÓN GENERAL DE CLASIFICACIÓN Y DATOS PERSONALES

REPORTE – REUNIÓN

DEPENDENCIA/ ENTIDAD: INSTITUTO FEDERAL DE ACCESO A LA INFORMACIÓN PÚBLICA (IFAI)
 ASISTENTES: Francisco Ciscomani Freaner- Secretario de Acuerdo- IFAI
 Lina Ornelas – Directora General de Clasificación y Datos Personales- IFAI
 LUGAR: Sala de Juntas del Pleno del IFAI
 FECHA: 24 de junio de 2005.
 ASUNTO: Abordar lo relativo al Recurso de Revocación OPINIO, en relación con la información de los gasoductos de PEMEX Gas y Petrosbras S.A.S.
 DESARROLLO: El Secretario de Acuerdo del IFAI manifestó la problemática existente en la determinación de la ubicación y de la información relativa a la ubicación de los gasoductos de PEMEX Gas y Petrosbras S.A.S. y los materiales con que son fabricados, entre los que destacan los siguientes:
 • Dentro de la cadena de Petrosbras, Petrosbras Gas ocupa una posición estratégica, al tener la responsabilidad del procesamiento del gas natural y sus líquidos, así como del transporte, comercialización y distribución de sus productos.
 • Petrosbras Gas es una de las principales empresas procesadoras de gas natural, con un volumen procesado durante 1999 de 3,527 millones de pies cúbicos diarios (mmpcd) y la segunda exportadora de líquidos, con una producción de 446 miles de barriles diarios (mmbd). Cuenta con una extensa red de gasoductos a través de la cual se transportan cerca de 100 mmpcd de gas natural, lo que la ubica en el 100 lugar entre las principales empresas exportadoras de este energético en Norteamérica.
 • En el año 2004 el Secretario de Acuerdo como la Directora General de Clasificación y Datos Personales señalan lo siguiente:
ELIMINADO: Un párrafo con tres regiones. Fundamento legal: Artículo 113 fracción VIII de la Ley General de Transparencia y Acceso a la Información Pública. En virtud de tratarse de información que contiene opiniones, recomendaciones o puntos de vista que forman parte de un proceso deliberativo de los servidores públicos, hasta en tanto no sea adoptada la decisión definitiva.
 • El sector energético, y en particular el de los hidrocarburos, ha sido una plataforma fundamental para el crecimiento económico de nuestro país. México no sólo cuenta con abundantes reservas de petróleo crudo y gas, sino que ha desarrollado una industria petrolera de gran complejidad y valor.
 Se acordó que se elaboraran diversos estudios para determinar la procedencia de la publicidad de la información señalada, toda vez que aún no se cuentan con elementos suficientes para emitir una opinión definitiva.
 ACUERDOS:
 (R.- 228912)

Ilustración 2 Modelo para testado de documentos electrónicos [Ilustración], por Consejo Nacional del Sistema Nacional de Transparencia, Acceso a la Información Pública y Protección de Datos Personales, 2016, Lineamientos generales en materia de clasificación y desclasificación de la información, así como para la elaboración de versiones públicas,
 (https://www.dof.gob.mx/nota_detalle.php?codigo=5433280&fecha=15/04/2016#gsc.tab=0)

De lo anterior es posible dar cuenta que el testado de las versiones públicas no permite el acceso completo a la información. Si bien se protegen los datos personales, el documento pierde parte de su información, por lo que el usuario podría tener dificultades para interpretarlo en la lectura. De esta forma, podemos intuir que la relación entre la privacidad y la transparencia es una relación tensa, dado el riesgo que se pone en juego en este procedimiento para la protección de datos personales.

El Archivo General de la Nación de Colombia y el Centro Nacional de Memoria Histórica(2022) definen la anonimización del siguiente modo:

Es un proceso técnico que hace que la información de identificación sobre las personas no sea identificable a quienes pertenece la información. Lo cual da como resultado el ocultamiento de

la información que identifique plena o parcialmente a las personas, organizaciones o características individuales de la fuente de información (2022)

En la ilustración 3 se puede representar como el testado que regularmente se realiza ofrece la máxima privacidad, pero a cambio da una utilidad mínima de la información. En el caso contrario un enfoque con privacidad mínima aumenta la utilidad de la información, pero pone en riesgo la privacidad de las personas. Hallar un equilibrio entre ambos extremos es la solución que más flexibilidad nos dará para el aprovechamiento de la información que se halla en los datos a la vez que se protege la privacidad de los implicados en los documentos.

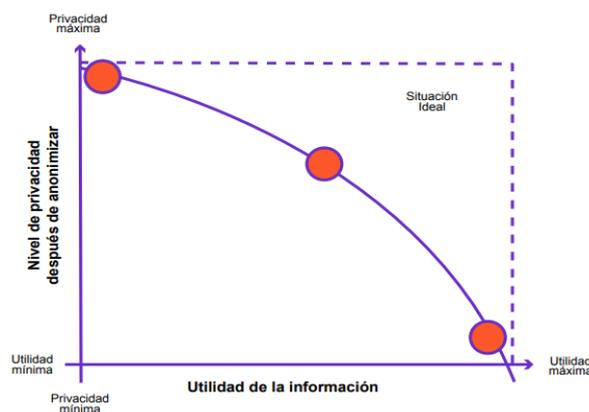


Ilustración 3 Adaptado de Luk Arbuckle, Privacy Analytics [Ilustración], por Archivo General de la Nación de Colombia, 2020, (https://www.archivogeneral.gov.co/sites/default/files/Estructura_Web/5_Consulte/Recursos/Publicaciones/Guia_de_Anonimizacion-min.pdf)

El mero hecho de no haber un ocultamiento o supresión total de los datos, como ocurre con el testado convencional, conlleva un riesgo de una posible identificación de las personas, principalmente por algún interesado que trate de inferir qué procedimiento llevó a la nueva presentación del documento. Tal como lo mencionan el Archivo General de la Nación (AGN) y el Centro Nacional de Memoria Histórica (2022): “la finalidad del proceso de

anonimización es eliminar o reducir al mínimo los riesgos de reidentificación de los datos anonimizados manteniendo la veracidad de los resultados del tratamiento de estos”.

Tal como lo indican Rangel et al. (2020, p.13), cuando los interesados no pueden ser identificados ni siquiera con los datos originales, los datos se anonimizan, ahí los documentos pueden utilizarse, reutilizarse y difundirse sin vulnerar los derechos de los implicados.

El proceso de anonimización puede ser resumido en los siguientes pasos (Centro Nacional de Memoria Histórica y Archivo General de la Nación, 2022, pp. 11-15):

- 1. Conformar un equipo de trabajo:** Se requiere formar un equipo interdisciplinario de trabajo, que integre la gestión documental y la tecnología.
- 2. Identificar qué tipo de datos requiere anonimizar:** Del tipo estructurados (tablas, bases de datos, hojas de cálculo) o no estructurados (texto, video, audio).
- 3. Identificar y clasificar los atributos:** Identificar y clasificar los atributos de los datos, incluyendo identificadores directos (número de seguridad social, CURP, nombre), identificadores indirectos o cuasi-identificadores (género, lugar de procedencia, fecha de nacimiento) y los identificadores directos e indirectos de tipo clasificado (orientación sexual, estado de salud).
- 4. Determinar y aplicar las técnicas de anonimización:** En análisis para elegir el uso correspondiente de la técnica adecuado de anonimización, según el caso.
- 5. Evaluar el riesgo e impacto de reidentificación:** Considerando el propósito de la publicación o intercambio de información, los niveles de acceso, la circulación interna y externa de la información, se realiza el cálculo del riesgo de reidentificación.
- 6. Evaluar la utilidad de los datos:** Considerando el equilibrio entre el nivel de privacidad o protección de los datos personales y su utilidad.
- 7. Publicar o compartir la información:** Procesos sistemáticos de publicación o intercambio que tengan medidas adecuadas de seguridad de la información.

- 8. Realizar un informe de anonimización:** Detallando los procedimientos, técnicas y metodologías usadas, incluyendo la trazabilidad de las gestiones, el análisis de riesgos y los responsables del proceso de anonimización.

Durante la evaluación del riesgo habrá que considerar si se trata de acceso a la información pública o de preparación de información para fines de análisis e investigación, por ejemplo, por lo que el riesgo es diferente si un grupo reducido de investigadores tiene acceso, en comparación con la población en general. Evaluar la utilidad de los datos está ligado también a la situación particular, por lo que se debe tener en cuenta, por ejemplo, si habrá fases posteriores de investigación para determinar la utilidad de los datos. A continuación se presenta la descripción realizada por Rangel et al. (2020, p.56-64). En la ilustración 4 se muestra un diagrama de flujo donde se muestra el proceso anterior.

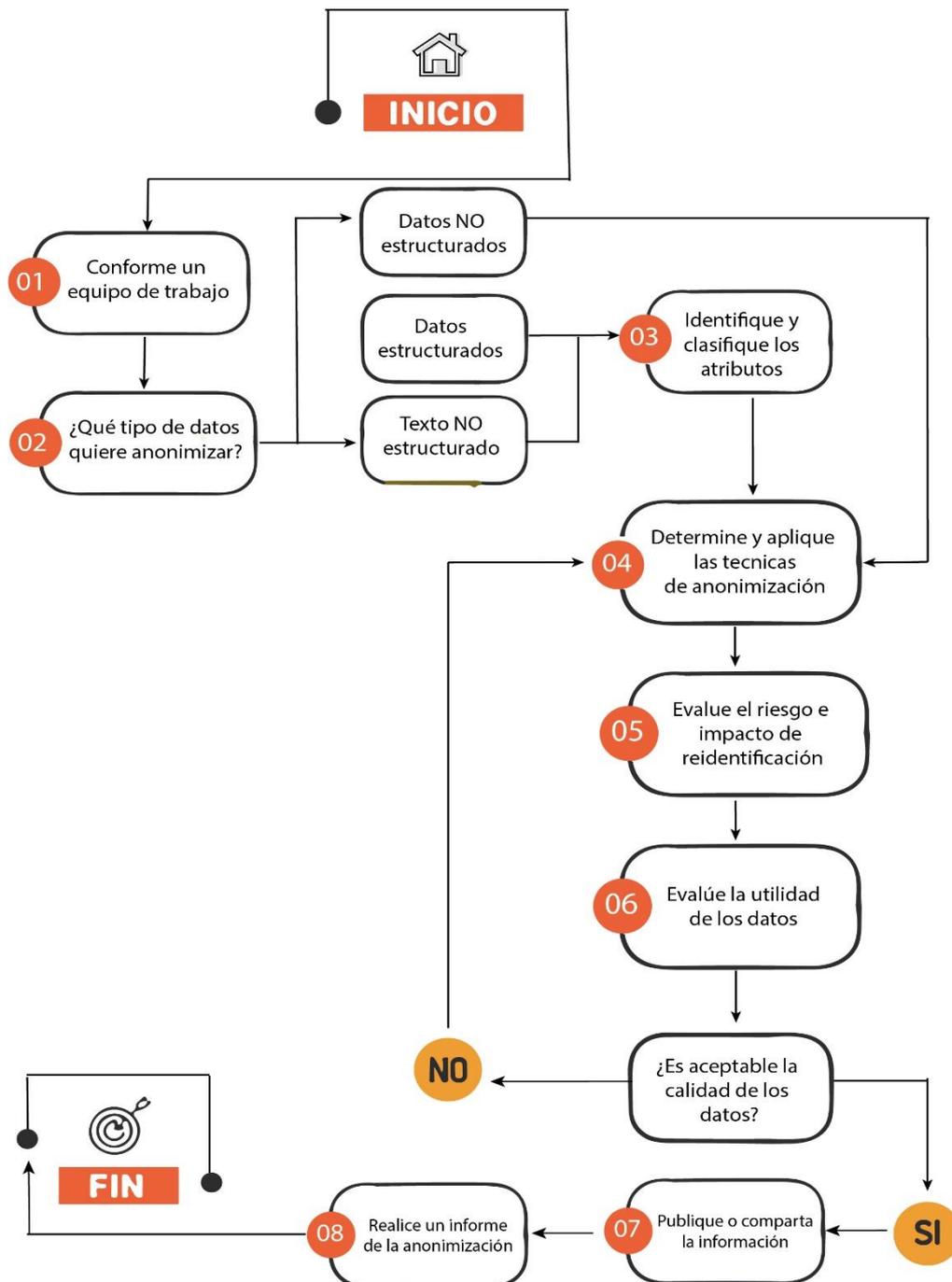


Ilustración 4 Centro Nacional de Memoria Histórica y Archivo General de la Nación. (2022). Diagrama para la anonimización de datos [Ilustración]. Guía para la anonimización de datos e información no estructurada: Estándares y lineamientos técnicos.

(<https://centrodememoriahistorica.gov.co/wp-content/uploads/2022/08/GUIA-DE-ANONIMIZACION.pdf>)

La anonimización no se limita a la aplicación de una técnica en específico, sino que conlleva trabajo interdisciplinario, análisis de la información, análisis de riesgos y documentación de los métodos utilizados. Para determinar la técnica de anonimización adecuada dependerá principalmente de factores como el tipo de identificador encontrado y del uso que se le va a dar a los datos. Entre las principales técnicas empleadas, se hallan la aleatorización, la generalización y la seudonimización

La aleatorización consiste en modificar los datos personales de modo que se rompa la vinculación entre datos que refieren a una persona. Puede ser por ejemplo un nombre asociado a una edad, cambiando la edad por otra diferente, ese cambio de valores se conoce como "ruido". Otro ejemplo es intercambiar datos pertenecientes de una persona y asignarlos a otra, técnica que se conoce como permutación.

La generalización permite ocultar detalles precisos sobre los datos personales al reemplazarlos por información de nivel más general. Por ejemplo, si tenemos los datos "Pedro Gómez nació el 12 de abril de 1985 en Bogotá", al aplicar generalización y se transforma a esto "Colombia en abril de 1985". Así ocultamos detalles específicos como el nombre de la persona, el día de su cumpleaños y la ciudad, dando como resultado información de un nivel más general y no específica. Dentro de este tipo de técnica esta la agregación, que consiste en generar rangos de datos en lugar de un solo dato en específico. Un ejemplo de agregación es reemplazar la edad de un individuo por un rango con varias edades. Por ejemplo, para una persona de 32 años, se genera el rango, de 30 a 39 años. Otra técnica es la supresión, que consiste en eliminar por completo algunos atributos o valores de los datos, esto se logra transformando el dato a asteriscos.

Hay otro proceso llamado desidentificación que es un caso particular del anterior, en términos de ser una técnica que no anonimiza de forma total, pero que permite aprovechar la información para su estudio o análisis. Lo anterior se logra al generalizar los datos

alejándose lo menos posible del valor real de los datos (Oficina del Comisionado de información y Privacidad de Ontario(2016, p. 5).

La técnica de seudonimización se basa en cambiar un atributo por otro dato, pero con la característica que en este caso los datos intercambiados identifican de forma única a una persona. Según este tipo de técnica se subdivide en: función hash, cifrado con clave secreta y descomposición en tokens (palabras). Las técnicas anteriores de forma general transforman los datos a un código alfanumérico o los encriptan; en ambos casos se requiere de la clave para desencriptar. Así, por ejemplo el nombre Miguel se transforma a E7889OP9QORBD098, en el caso de usar la función hash, por dar un ejemplo.

Entre los riesgos que puede tener la anonimización se hallan los siguientes (Barranquero, 2022, p.11):

1. **Singularización:** Se basa en la posibilidad de identificar una persona o varias mediante el uso de varios atributos.
2. **Vinculabilidad:** Se basa en hallar una relación entre dos o más atributos que identifiquen a una persona. Para el caso de un grupo de personas, es posible hallar varios atributos que las asocien, pero si no es posible identificar a una de ellas, se puede decir que si bien esta el riesgo presente de la vinculabilidad, es resistente a la singularización.
3. **Inferencia:** La posibilidad de deducir un atributo en un conjunto de datos.

La anonimización no es un proceso que resulte carente de riesgos. La seudonimización es débil en la singularización ya que es posible identificar a un individuo, pero es fuerte ante la inferencia porque está encriptada con códigos alfanuméricos. La aleatorización es fuerte ante la vinculabilidad ya que los datos son intercambiados o convertidos en rangos, pero débil ante la singularización, si los atributos de una persona son dejados en una misma fila. Para remediar esto pueden ser también permutados a la vez que se añade el ruido. Por último, la generalización es fuerte ante la singularización al generar

un rango que impida precisar los atributos de una persona en específico, pero es más débil frente a la vinculabilidad y la inferencia.

Así pues, cada técnica de anonimización no siempre se aplica de forma aislada, se debe tener el criterio suficiente para establecer la combinación de técnicas adecuadas según el caso. La anonimización en general es un proceso que debe llevarse a cabo con un análisis adecuado para lograr los mejores resultados, considerando tanto los fines del aprovechamiento de la información como el logro de un balance entre la transparencia y la privacidad, entendiendo, que, entre más se consiga, de mayor beneficio será el acceso a la información

La inteligencia artificial y la anonimización

Como ya se mencionó, la anonimización es una técnica que requiere un cuidadoso análisis en su aplicación. Regularmente, cuando se aplica a datos estructurados como bases de datos, hojas de cálculo y formularios, es sencillo identificar los atributos o columnas que contienen la información de interés. Por otra parte, para el caso de datos no estructurados, como texto libre, la identificación no es sencilla porque no existe un identificador de columna asociado, por lo que se requerirá de recorrer toda la información. Además, en este caso, como los datos se leen en narrativa, es decir, que en la redacción hay información que da contexto y sentido a los datos, aumenta la probabilidad de encontrar elementos diversos que permitan vincular los datos con las personas. Esto incrementa la dificultad de la tarea.

Es razonable pensar que la IA es apropiada para la anonimización de datos personales, cuando se trabaja con grandes volúmenes de datos. La incorporación de esta tecnología al flujo de trabajo podrá aumentar la velocidad en el tratamiento de los datos, pero también requerirá analizar su fiabilidad para ayudar en el caso de la anonimización. En particular, el uso de la IA para procesar grandes cantidades de datos conlleva el riesgo de no cumplir con la legislación aplicable en protección de datos personales, de modo que el ser capaz de identificar y clasificar correctamente la información sensible y de identificación

personal es un primer paso para el uso de la IA para brindar mayor acceso a los datos personales.

Jaillant y Caputo (2022, p. 831-832) examinan los retos en materia de protección de información sensible utilizando métodos de IA como un mecanismo de transparencia. Para las autoras, la información personal se refiere a cualquier información que pueda identificar de forma exclusiva a un individuo, desde nombres, números de teléfono y de seguridad, hasta datos de nacimiento y médicos, incluyendo también información clasificada y legal, así como datos resultantes de experimentos de investigación. En la vinculación del uso de métodos de IA y la protección de información sensible, toman un enfoque que se basa en la identificación y cuantificación de la información sensible.

La identificación puede ser abordada con algoritmos de clasificación de textos, en los que se utiliza un conjunto de rasgos característicos extraídos de los documentos para predecir las clases o categorías de un documento. En la clasificación, las características suelen extraerse directamente del texto en forma de palabras clave, es decir, los términos individuales que constituyen el texto. Esta técnica ayuda a clasificar documentos como artículos científicos, noticias y correos electrónicos.

La revisión de sensibilidad o identificación de datos personales sensibles en el texto no se orienta específicamente a identificar términos específicos, sino más bien identificar quién dijo qué sobre quién, pues en las relaciones entre los datos no necesariamente personales y las personas mencionadas en el texto puede inferirse información sensible; así, los esfuerzos se centran en identificar relaciones entre palabras y entidades en el texto, en lugar de palabras clave específicas. Encontrar una conexión entre el nombre de una persona y una actividad ilegal podría implicar que un dato sea funcional para re-identificar a una persona. Deberían anonimizarse los datos que revelen esa relación.

Existe la idea errónea de que el estándar que se debe alcanzar es el realizado por las personas. Si consideramos que la cantidad de información a revisar es inmensa y pueden existir errores y sesgos cognitivos en las personas, un mejor enfoque sería usar la IA como herramienta complementaria. Otro enfoque alternativo es incorporar a las personas dentro

del ciclo de desarrollo, por lo que se piensa en las soluciones de AA no como productos únicos, sino como asistentes digitales vivos que aprenden junto a los revisores humanos.

Esta modalidad se adapta a casos donde la sensibilidad adecuada no se conoce a priori; a esto se le conoce como Tecnología de Revisión Asistida. El flujo de trabajo se basaría en tener desde un inicio un conjunto de documentos ya revisados, destinado a que la IA genere un nuevo conjunto de documentos a revisar por personas y, posteriormente, este nuevo conjunto amplíe el nivel de entrenamiento de la IA. Se generan de esta forma ciclos de entrenamiento con el fin en colecciones nuevas de documentos.

Las autoras concluyen que el uso de la IA debe regirse por un marco de gobernanza y unos procedimientos bien desarrollados en materia de consentimiento, inclusión, transparencia, ética y privacidad, motivados por las prácticas establecidas entre los archivistas, así como de los científicos sociales, los historiadores y los antropólogos.

A partir de las ideas de Jaillant y Caputo, cabe puntualizar que, en el esquema de incluir personas en el ciclo, como archivistas somos capaces de identificar contextos y contenidos particularmente diferentes que les permitan decidir que documentos podrían ser suministrados a la IA o no, a la vez que son conscientes de manejar el posible margen de error. El uso de esta tecnología, cuanto más cuando incide directamente en decisiones importantes para las personas, requiere un cuidadoso análisis para evitar una reproducción sistemática de errores en el procedimiento que resulten en un trato; esa es otra de las bondades del esquema propuesto.

A continuación, se presentan los hallazgos en la revisión de literatura sobre estudios de caso de aplicaciones de la IA para la anonimización en las áreas médicas y legales. Cabe señalar que se han seleccionado los artículos que resultaron de mayor interés para el proyecto de investigación sobre ciencia de datos en salud sobre la diabetes, citado más adelante, que sirvió como marco para la realización de este trabajo.

Las investigaciones sobre salud a gran escala, que requieren el análisis de documentos, necesitan de datos anonimizados, de forma que no se invada la privacidad de

las personas, especialmente en tanto que los datos sobre la salud son datos personales sensibles. A partir de la revisión realizada por (Meystre et al., 2010) sobre el uso del aprendizaje automático para la anonimización de documentos generados en el área médica, se reporta que en el caso de documentos de textos no estructurados como notas de paciente, el aprendizaje automático brinda ventajas en la anonimización de datos personales. Se adapta de forma autónoma a los diferentes datos suministrados, pero tiene como desventaja el hecho de no poder precisar las reglas de identificación de forma específica, si no que requieren de suministrar conjuntos amplios de datos para el entrenamiento que previamente hayan sido etiquetados (Meystre et al., 2010, p. 7).

Entre los estudios hechos para la anonimización se halla el de Gardner y Xiong, donde se utilizó aprendizaje automático junto con otro modelo estadístico, para la identificación y anonimización de datos personales y datos personales sensibles en 100 reportes de patología del Winship Cancer Institute. Los documentos fueron etiquetados manualmente respecto a los siguientes datos personales: nombre, fecha de nacimiento, edad, número de documento médico, y números de cuenta. Se utilizó una pequeña parte para entrenar a la IA que posteriormente se usó para analizar los reportes restantes. Se verificaron los resultados y se corrigieron los reportes que fueron mal identificados. En el reconocimiento de datos personales los resultados tuvieron una alta calificación respecto a la fecha, mientras que el dato con la menor efectividad correspondió al nombre (Gardner & Xiong, 2008, pp.9-10). En el caso anterior es posible percatarse que datos como la fecha son fácilmente identificables cuando son registradas en un formato más estandarizado, en tanto que la fecha resulta una combinación de números con una estructura regular, mientras que la diversidad de nombres es muy amplia y dificulta la tarea.

Otro caso de la misma área para el caso de México es el indicado por el Instituto Mexicano del Seguro Social (IMSS), donde en el estado de Michoacán se concretó un estudio de anonimización sobre notas clínicas para la investigación en la predicción de la diabetes mellitus tipo 2. Esto fue necesario ya que, aunque la base de datos ya había tenido una anonimización previa como conjunto estructurado, se identificó que los médicos, al llevar

a cabo los registros de consultas, anotaban los datos personales del paciente como nombre, dirección o teléfono. Para efectuar este estudio se hizo uso de herramientas de PLN que habían sido específicamente entrenadas en español, pero que contaron con una segunda fase de entrenamiento específicamente con narrativas clínicas (Álvarez Aguilar et al., 2022, p. 12). En este caso es interesante notar que, según la fuente de los datos y el propósito de la investigación, fue necesario hacer ajustes particulares para adoptar técnicas computacionales de uso más general.

De acuerdo con los tres estudios, el uso de modelos de aprendizaje automático para anonimizar documentos clínicos con datos no estructurados presenta un buen rendimiento, en especial cuando se implementan en un tipo específico de documento, pero representa un desafío generalizar su aplicación a una amplia variedad de documentos, sistemas y datos personales correspondientes. Otra cuestión es como el uso de la anonimización afecta la lectura e interpretación del texto (Meystre et al., 2010, pp.14-15).

Esta última preocupación es compartida también en los casos del área legal, los documentos médicos hacen regularmente referencia a una sola persona, es decir el paciente. Por tanto, es usual que los datos personales o sensibles se relacionen a una única persona; mientras que en el área legal se puede hacer referencia a las personas involucradas en un mismo documento, por lo que se puede inferir es mayor. Con esto en mente, podemos intuir que el testado de documentos legales dificulta más la lectura que el que se realiza sobre expedientes médicos.

Un caso de anonimización es el presentado en Uruguay a partir de la Base de Jurisprudencia Nacional Pública (BJNP), que brinda acceso público y gratuito a las decisiones judiciales desde 2008. En las versiones publicadas, se ocultan datos personales sensibles para evitar consecuencias dañinas, como invasión de privacidad. Entre los datos ocultados están los nombres de los infractores, testigos y acciones criminales, por mencionar algunos. El uso de la IA en este proceso se ha propuesto para reducir carga y tiempos de trabajo, mejorando la experiencia de usuario (Garat & Wonsever, 2022, pp. 1-2).

Este caso buscaba un objetivo específico más allá de la anonimización: mantener la referencia de las personas en el texto. Si se reemplazan los nombres por otros al azar o se enmascaran con asteriscos, se perdería la coherencia durante la lectura del texto, ya que al leerse no se podría identificar con certeza a qué persona se refiere en cada momento o cuántas personas son nombradas. Así, se buscaba no solo identificar los nombres sino también lograra mantener la referencia a la persona que aludía el texto, lo que se conoce como co-referencia; por ejemplo, para cada persona se utilizó un seudónimo, que será único para cada persona referenciada en el texto. Por ejemplo, "Sr. Miguel Fuentes" y "Sr. Fuentes" se reemplazan por "Sr. AA", porque ambos se refieren a la misma persona, mientras que "María Rodríguez" y "Juana Fernández" se reemplazan por "BB" y "CC" respectivamente (Garat & Wonsever 2022, pp. 5-6).

Para este caso se contaban con 80,000 documentos, 17,000 estaban des-identificados. Durante la revisión de estos documentos se notaron inconsistencias en el proceso de des-identificación, así que fue necesario identificar y etiquetar manualmente 1000 documentos. De ese lote, se seleccionaron para el entrenamiento 800 documentos, mientras que los 200 restantes para la validación del modelo. En la métrica de evaluación se encontraron resultados satisfactorios no únicamente para la identificación de datos, sino también para la corrección en la co-referencia al anonimizar (Garat & Wonsever, 2022, p. 6).

Los autores señalan que, dado el riesgo en particular en relación a la protección de datos personales, el procedimiento de anonimización no puede automatizarse en su totalidad. Un mejor modo de integrar esta herramienta al proceso de anonimización es como un asistente del personal que realiza el proceso. Al mismo tiempo, el uso de esta herramienta puede servir para retroalimentar al modelo con futuros documentos que sean correctamente anonimizados con ayuda de personas específicas y la IA (Garat & Wonsever, 2022, pp. 12-13). En este caso específico, los efectos particulares de la anonimización en la comprensión de su contenido delimitaron las métricas de validación del algoritmo; así la consideración de la experiencia del usuario fue determinante para plantear los objetivos del modelo e incluir la co-referencia.

A partir de un análisis teórico sobre la anonimización automatizada en el contexto del sistema legal húngaro, los investigadores Csányi et al. (2021) proponen no sólo utilizar el Reconocimiento de Entidades Nombradas (REN), sino también el Reconocimiento de Entidades Enlazadas, (REE) para así establecer relaciones entre datos como el nombre de la persona, la institución y la profesión. Por ejemplo, que permitan identificar riesgos de vinculabilidad, inferencia o singularización; además, sugieren incluir el Reconocimiento de Eventos (RE) para identificar palabras clave que permitan reconocer hechos. (Csányi et al., 2021, p. 20).

Enfoques como el anterior permitirán ampliar los criterios para evaluar la robustez de una solución de anonimización, considerando no únicamente la identificación de datos sensibles, sino también una correcta aplicación de la técnica elegida. Es importante tener claro que agregar más componentes al procesamiento no significa necesariamente que sea mejor en términos de calidad, por lo que será un análisis interdisciplinario para determinar qué solución ofrece suficiente rendimiento sin tener que asumir una complejidad innecesaria.

Analizar qué datos del documento podrían presentar un riesgo para la re-identificación según la técnica de anonimización es uno de los procesos donde el archivista puede presentarse como un participante activo. Es posible ayudar como archivistas en la revisión de tales documentos para el reconocimiento de ajustes precisos que mejoren la identificación, así como señalar las dificultades en la lectura del documento final y reconocer riesgos de re-identificación.

[La fiabilidad de la inteligencia artificial en la anonimización de datos personales](#)

A lo largo del presente trabajo se ha dado lectura a textos sobre el uso de la IA en diferentes casos de estudio, que ayude para ubicar nuestro papel como archivistas y las particularidades del funcionamiento de la IA que sugieran como la archivística podría integrar esta tecnología en sus funciones. En general, hemos visto que la IA se utiliza para procesar grandes volúmenes de información, en varios casos como documentos digitales, pero también como objetos computacionales susceptibles de tratamientos archivísticos, como

registros contenidos en bases de datos. El tiempo actual donde la archivística entra en contacto con la IA sitúa grandes retos hacia como incorporamos nuestra práctica profesional en un espacio más centrado en los datos.

Como lo mencionan Marciano et al (2018, pp.179-199), la producción y el consumo de los documentos digitales está influido por las tendencias sociales e industriales junto con la computación, que no muestran conexión directa con la archivística. Deberíamos adaptarnos como archivistas a este nuevo entorno del procesamiento de la información y tratar de comprender estos nuevos métodos, entre los cuales se incluye la IA, a la vez que se informa a la tecnología de cómo deberían ser tratados los documentos y los datos, para su preservación y acceso.

Este contacto nos acerca a un área de conocimiento propuesta recientemente, la Archivística Computacional (CAS), por sus siglas en inglés), que Marciano et al (2018) la definen como:

Ámbito interdisciplinario relacionado con la aplicación de métodos y recursos informáticos al tratamiento, análisis, almacenamiento, conservación a largo plazo y acceso a gran escala de documentos/archivos, con el objetivo de mejorar la eficacia, productividad y precisión en apoyo de la valoración, ordenación y descripción, las decisiones de conservación y acceso, y la participación y realización de investigaciones con material de archivo.

El proyecto en el que está inscrito este trabajo se puede enmarcar en este campo donde, a partir de la archivística, se busca la fiabilidad del uso de esta tecnología, lo que requiere un involucramiento más profundo de ambas disciplinas con una mayor comprensión del funcionamiento de la IA. El diálogo entre la archivística y las ciencias de la computación permitirá también profundizar en el tratamiento requerido por los documentos, vistos como datos, para el entrenamiento de la IA y el procesado de grandes volúmenes de información; como ejemplo concreto, vemos que el archivista puede brindar contexto, al dar los requerimientos necesarios para el procesamiento específico de la información que custodia, lo cual es esencial para la eficacia, la precisión y la pertinencia de un modelo de IA.

Cada vez resulta más necesario el análisis de la fiabilidad en la IA, ya que como se ha visto anteriormente su uso sin un análisis cuidadoso y crítico puede llevar a cometer injusticias en forma de sesgos, discriminaciones e injusticias. El tema de la fiabilidad en la IA es una cuestión que recientemente se ha venido estudiando; se comentarán a continuación algunas perspectivas.

Para Harrison y Luna-Reyes (2020, p. 501) entre los elementos que deberían ser analizados para una IA confiable se halla la confiabilidad de los datos. Se requieren datos de calidad que se consiguen al haber una correcta gestión de datos durante la fase de recogida; además, al tener un proceso documentado podrá haber una replicación de los resultados de la IA. Adicionalmente, entre las múltiples fuentes de datos de las que la IA se alimenta, es probable que algunas de ellas tengan datos privados, porque es necesaria la seguridad en el uso de tales datos, para evitar riesgos de re-identificación. Como solución a los dos problemas anteriores se propone:

1. Realizar una gestión de datos basada en la gobernanza, donde se regule como su tratamiento mediante el uso de algún marco de trabajo que brinde directrices para la correcta gestión de datos en áreas desde la arquitectura y diseño hasta la gestión de metadatos. Mediante esto es posible minimizar los problemas de seguridad al establecer medidas básicas para la creación, uso y acceso.
2. Capacitar sobre gestión de datos al personal encargado del control, para que tenga los elementos necesarios para realizar su labor.

Cabe señalar que nosotros los archivistas somos un importante referente para formar parte de los dos procesos propuestos por los autores, ya que somos especialistas en procesos de gestión específicos para archivos y colecciones que contienen los datos.

En la literatura aparece una tensión entre la vulnerabilidad y la confiabilidad en las herramientas computacionales. Por una parte, Harrison y Luna-Reyes (2020) plantean que, en la medida en que los modelos sean más explicables y transparentes, puede aumentar el riesgo de inferir la información anonimizada. Por otra parte, Hamon et al., hacen ver que las

nociones técnicas de explicabilidad y transparencia se convierten en requisitos cruciales para garantizar que los sistemas de toma de decisiones basados en la IA aplican de forma adecuada los marcos regulatorios de protección de información (Hamon et al., 2022, pp.74-75). En el caso de México, la GDPR nos brinda directrices de lo que podría esperarse del uso de la IA en el contexto de los datos personales, para los que sería necesario realizar un estudio específico.

La transparencia de un sistema va más allá de sólo conocer los parámetros del sistema sino es tener un conocimiento suficiente de todo el proceso de desarrollo, es decir, incluyendo las fases de procesamiento de datos, entrenamiento y evaluación (Hamon et al., 2022, p.75.). Los casos de estudio mencionados a lo largo de la investigación muestran un proceso integral interdisciplinario donde se realiza no únicamente el entrenamiento del algoritmo, sino también entrevistas, descripción de los documentos utilizados y evaluaciones necesarias para la mejora del algoritmo, que deberían a su vez transparentarse.

Hamon et al. (2022) proponen que, en la búsqueda de una transparencia o rendición de cuentas, se debe pensar más allá de la explicación técnica. Por lo que proponen el uso de una Evaluación de impacto de protección de datos (DPIA), donde el responsable de los datos debería describir el tratamiento de datos, evaluar la necesidad y proporcionalidad del mismo, analizar los riesgos para los interesados (por ejemplo, riesgos de discriminación, violación del derecho a la salud, a la libertad de pensamiento, de movimiento, ataque a su identidad digital, etc.), y solicitar la opinión de los interesados sobre el tratamiento de datos (Hamon et al., 2022, pp .82-83). Cabe señalar que el archivista puede en este contexto fungir como controlador de datos ya que sus funciones están alineadas con las nuevas necesidades en términos de responsabilidad en el manejo de los datos.

Para Hamon et al. (2022 pp. 82-83) la robustez de los sistemas de IA está interrelacionada con la explicabilidad. Un modelo robusto, comprobado sistemáticamente y de forma adecuada dará lugar a resultados más sólidos y fiables, aunque no sea directamente legible para el ser humano en esta lógica, la decisión sobre el uso de

herramientas de evaluación y controles que resulte en un algoritmo de mayor exactitud, deberá ser debidamente documentada, lo que daría como resultado una explicabilidad por diseño.

Como conclusión, (Hamon et al., 2022) proponen que un conjunto completo de métricas para una IA puede ser más efectivo que mitigar riesgos de un componente específico; es decir, que se debe tener una visión general en cómo los componentes del sistema y su proceso de desarrollo dan forma al diseño de los sistemas de IA los cuales no sean injustos, ilegales o inexactos, preferible que una visión que se preocupa únicamente por los principios de protección de datos. El enfoque propuesto es que, en vez de aplicar un estándar a cada principio de protección de información, sean sopesados los riesgos en conjunto (Hamon, 2022, p.84). Así, los principios se interrelacionan por lo que el diseño a su vez se vuelve más robusto en el sentido de tomar en cuenta varios factores y generar medidas para la cohesión de estos.

A lo largo de este informe se han retomado elementos de la revisión de la literatura que permitan delimitar un marco de acción donde, como archivistas podemos relacionarnos con la IA. Durante la revisión de la literatura, dentro de varias fuentes se identificó que el acercamiento a la IA desde los archivos se realizaba con énfasis en la protección de datos personales, en particular al GDPR y la privacidad, pero también en función del procesamiento de datos para un modelo de IA con mayor fiabilidad y responsabilidad (Geburu, 2020, Andresen, 2019, Jailliant & Caputo, 2022).

Asegurar la fiabilidad de la IA en la protección de datos personales parte de comprender los elementos que la conforman y su proceso de desarrollo e implementación. La IA en sí misma no es un agente o entidad independiente, los resultados generados por la IA son el reflejo del desarrollo que llevó a su creación; además, este proceso no debería ser entendido como un proceso finito y aislado de aquellas personas no expertas en desarrollo de IA. La literatura demuestra que el grado de apertura durante su diseño e implementación, que implique el seguimiento de normas aplicables junto a la colaboración activa de los

responsables donde se hará uso de esta tecnología, sumado a un ciclo permanente de refinamiento de la IA, constituyen un mínimo de fiabilidad que puede ser alcanzado.

Tal como se indicó en la Tabla 6, como archivistas podemos cumplir varias funciones en relación a la IA; a partir de estas, es posible abordar el caso específico de la IA en la protección de datos personales.

A partir de lo anterior se mencionarán algunas acciones específicas en la siguiente Tabla 7, donde es posible formar parte como archivistas del proceso del desarrollo de la IA, recordando que puede haber otros futuros casos conforme la investigación en el área siga avanzando. Basado en lo anterior es posible enmarcar la fiabilidad en las distintas actividades que siendo archivistas podemos ejercer desde la profesión archivística, en el marco de las funciones de la archivística, ya que están se orientan principalmente con una orientación hacia a la transparencia y explicabilidad de la IA

Funciones	Actividades
Documentación	<ul style="list-style-type: none"> • Identificar y documentar la robustez de la IA en términos de su proceso de desarrollo expresado en el grado de cumplimiento de normativas, gobernanzas u otra clase de comprobaciones aplicables (Hamon et al., 2022).
Participación en el desarrollo de modelos	<ul style="list-style-type: none"> • Implementar una evaluación de impacto de protección de datos (Hamon et al., 2022). • Identificar y analizar de falsos positivos y falsos negativos (Hutchinson, 2018). • Participar en la valoración del grado de fiabilidad alcanzable mediante el análisis del equilibrio en la interrelación de elementos tales como: precisión y el rendimiento, la explicabilidad y la transparencia, así como la importancia mutua de la robustez, la equidad y la explicabilidad (Hamon et al., 2022). • Formular una política de valoración de los conjuntos de datos a utilizar para el entrenamiento de la IA vigilando la autenticidad, privacidad e inclusión (Gebru, 2020). • Promover códigos de ética (Gebru, 2020). • Realizar evaluación de riesgos y evaluación ética (Andresen, 2019).
Configuración	<ul style="list-style-type: none"> • Establecer pruebas de entrenamiento para verificar el rendimiento de la IA (Hutchinson, 2018). • Analizar cómo se comporta la IA con base a distintos conjuntos de entrenamiento (Hutchinson, 2018). • Si la IA está implementada en un software de aplicación, realizar diferentes pruebas de configuración para analizar el impacto en la actividad a realizar (Hutchinson, 2018).
Análisis crítico de los datos	<ul style="list-style-type: none"> • Tratar de identificar elementos diplomáticos o archivísticos que ayuden al análisis crítico de las limitaciones del uso del algoritmo y en la gestión a partir de las especificidades de sus datos de entrenamiento (Chabin, 2020; Gebru, 2020).
Ética profesional	<ul style="list-style-type: none"> • Identificar la interpretabilidad del modelo de aprendizaje y participar en su comprensión (Hamon et al., 2022). • Colaborar en establecer una justificación del uso de la IA que demuestre no es injusta, ilegal o inexacta. (Hamon et al., 2022).
Pertinencia y mejora	<ul style="list-style-type: none"> • Realizar recomendaciones a partir de la experiencia de usuario o el conocimiento especializado sobre los conjuntos documentales (Garat & Wonsever, 2022).

Tabla 8. Funciones y acciones del archivista en relación a la fiabilidad de la IA en la protección de datos personales. Fuente: Elaboración propia

Consideraciones finales

La IA abre un panorama tecnológico que puede ser un buen complemento para las distintas funciones dentro de un archivo. Para aquellos sin contacto previo con ella, su funcionamiento puede resultar confuso, ya que requiere conocimientos especializados de matemáticas, estadística y computación. A nivel general es posible acercarse a conceptos relacionados con la IA como: los tipos de aprendizaje, las redes neuronales, el procesamiento del lenguaje natural y sus distintas aplicaciones, además del flujo de desarrollo de la IA y los posibles riesgos. Dentro de estos conceptos aparecen otros derivados según las necesidades del caso en específico que impliquen desarrollar tecnologías adicionales.

La comprensión del desarrollo de IA, a nivel superficial, puede resultar sencilla de bosquejar: comienza por tener un conjunto de datos, para elegir un modelo matemático, estadístico y computacional, que detecta patrones cuya utilidad sea evaluada, volviendo a repetir el proceso hasta alcanzar un grado de estabilidad y eficacia. En miras al impacto que estos métodos pueden tener, se debe vigilar que el ciclo descrito se realice de forma adecuada, es decir, que cada implementación se adecue al contexto específico donde planea usarse.

Como archivista es posible cumplir con varias funciones durante el proceso de desarrollo de la IA, como ya se ha mencionado anteriormente aunque sea una propuesta. Hay aspectos en los que se puede tener una participación comprometida, y aprender de primera mano cómo la IA creará al documento o auxiliará en otras funciones. La interacción con el equipo de desarrollo es recomendable y necesaria, donde puede además asumir un rol protagónico para el desarrollo y seguimiento de principios éticos y de evaluación continua.

En un futuro queda por estudiar cómo integrar los instrumentos archivísticos, como inventarios, catálogos, guías, fichas de descripción, metadatos, de forma que podamos aportar más a este control de los conjuntos de datos. Aun así, la implementación requiere

de un proceso medido y controlado, donde puedan probar en primera instancia prototipos de estos modelos de IA, y no hacer uso de ellas directamente sin antes medir de forma directa su alcance. De la interacción en el proceso de desarrollo es posible inferir que existe un mayor control cuando éste incide en la producción del documento, siendo este el medio por el cual puede conocerse el grado de fiabilidad.

El inicio de la regulación de la IA se enfrenta a cierto grado de autonomía en el procesamiento de los datos, poco antes explorado. Incluso los mismos desarrolladores de la IA no pueden tener un control preciso de los resultados que genera, si hablamos en términos estadísticos, pese a que se lleguen alcanzar una precisión de más del 90%: dada la cantidad de datos que se manejan, el margen de error restante puede implicar que miles de datos sean erróneamente procesados. También cabe mencionar aquí que el que una persona realice el mismo procesamiento, no la exenta de errores, así que ambas partes, tanto algoritmos como personas, pueden complementarse.

Las estrategias de mitigación de sesgos y reducción de riesgos comienzan durante el desarrollo de la IA, analizando los datos utilizados para el entrenamiento en términos de cómo fueron recopilados y documentando las decisiones llevadas a cabo durante el desarrollo. La privacidad es otro componente que resulta fundamental en la protección de los usuarios del sistema, ya que la IA en algunos casos tratará con datos personales para su entrenamiento y para la realización de sus tareas.

Trasladar los principios a la realidad práctica es una tarea difícil pero, como se ha visto, al instrumentarse regularmente se expresan en términos del control en el desarrollo de un modelo de IA y su documentación que permita trazar las decisiones técnicas con posible implicación ética a lo largo del proceso. La efectividad en la aplicación de los principios y medidas prácticas tiene un componente altamente ético, se debe promover una conciencia acerca de los riesgos que un desarrollo acelerado y poco controlado podría generar. Como se sugiere en algunos de los textos revisados en la sección anterior, la archivística y sus instrumentos administrativos pueden añadir formas de gestión sobre los datos que se

trasladan desde los documentos, pero también sobre la creación y uso de conjuntos de entrenamiento en general.

En el caso de su uso para la protección de datos personales a nivel general, es posible afirmar que, al menos en lo que respecta a la identificación de los mismos, hay un alto grado de precisión. En cuanto a una anonimización que esté más allá de la supresión de la información o el simple intercambio por seudónimos, sin tomar en cuenta la coherencia del texto, se requiere aún de una mayor investigación y desarrollo. Para cada caso donde se implemente esta tecnología los componentes serán diferentes, esto debido a las diferentes requerimientos que pueda tener cada proyecto, en relación a las diferentes técnicas de anonimización; así, realizar una afirmación a nivel general resulta difícil. Lo que se puede asegurar es que en casos donde el documento anonimizado requiere de un tratamiento más especializado, más allá de la identificación de datos personales. Las relaciones de personas, eventos, etiquetado por co-referencia, son acciones que entre más específicas sean, la IA puede ser menos precisa. Por ello en esas tareas debería haber revisores que de forma manual o asistida con IA deberían realizarlas.

La IA nos brinda una gran oportunidad para brindar acceso a la información a grandes volúmenes de información, que de otro modo requeriría recursos para anonimizar de forma manual y sencilla. En el presente trabajo no se tocó el estado actual de los archivos en México, el cual es crítico, ya que en muchos casos no se tiene la infraestructura suficiente - tanto financieros como de espacio-, que permita una adecuada conservación de los documentos: mucho menos su digitalización si la condición anterior no es cumplida. Frente a esta problemática es posible decir que esta tecnología acentúa el estado actual de rezago que tienen los archivos en México y como esto obstaculiza la implementación de estas tecnologías en México y como esto obstaculiza la implementación de estas tecnologías.

Este panorama nos plantea dudas acerca del desarrollo que tendrá esta tecnología en el área de los archivos en nuestro país, sin embargo, esta puede ser una oportunidad para sensibilizar de forma especial considerando que las condiciones adecuadas los

archivos pueden servirse de la IA para dar el tratamiento de información necesario. Para ello, las condiciones de infraestructura, tecnología, capacitación y financiación deben ser cumplidas. La IA nos permite vislumbrar escenarios futuros, nos brinda aún más razones por las cuales es benéfico para la sociedad y el gobierno invertir en los archivos, en tanto una correcta organización y conservación de los documentos, junto con la infraestructura necesaria, establecería las condiciones para el uso de la IA lo cual amplificaría la capacidad de procesar la información y aplicar distintos tratamientos, como en el caso de la anonimización mencionado en este trabajo.

La hipótesis inicial de este trabajo era hallar claves relevantes entre la archivística y la IA para la identificación y extracción de datos personales y sensibles. Partiendo de la escasez de literatura que conjuntara los aspectos anteriores se conformó un marco contextual que sí nos ha permitido proyectar líneas de acción en la intersección entre IA, archivos y protección de datos personales. Sobre este hecho en particular, las funciones mencionadas en la Tabla 7 permiten que sea posible tener una base de interacción inicial en proyectos con datos personales e IA. Sobre las acciones específicas que puedan realizarse, como se ha visto, dependerá de las particularidades donde se implemente. La rigurosidad y exactitud del proceso de anonimización está determinada por los objetivos particulares en cada proyecto.

Hay elementos claves particulares mencionados en la Tabla 8, como la evaluación de impacto de protección de datos, la valoración del rendimiento de la IA tomando en cuenta la proporción de falsos positivos y negativos, el riesgo que el uso de la IA puede introducir en los procesos donde se incorpora, son los más destacables.

La anonimización en la protección de datos personales, requiere de la identificación y extracción de datos, pero la dimensión que puede adquirir esto va desde el simple reconocimiento de datos aislados, hasta el reconocimiento de eventos o la etiquetación de entidades nombradas para mantener la co-referencias. A su vez, el entrenamiento es una cuestión clave, dado que en determinados contextos quizá no se cuenta con la cantidad

necesaria de documentos para entrenar a la IA. Así que entre la definición del diseño de una solución de anonimización más o menos especializada y la fuente de datos que se pretende utilizar para conseguirlo, se debe realizar un estudio de riesgo, precisión, rendimiento, posibilidades de explicabilidad y necesidades de transparencia, como sugiere la Tabla 8.

Se puede decir que hay ciertas claves referentes al aspecto técnico, pero las más relevantes desde un punto de vista de uno como archivista incluyen la supervisión de un desarrollo ético y responsable, donde la implementación de la IA se realice teniendo como meta un aprovechamiento transparente y lo suficientemente preciso. Aquellos pasos de la anonimización con la IA que no alcancen un mínimo de rendimiento seguirán a la espera de mejores desarrollos de la IA, donde será indispensable contar con la intervención humana de especialistas en los conjuntos documentales utilizados.

En síntesis, se debe abogar por una aplicación de esta tecnología de forma regulada, medida y segura, evitando un proceso acelerado de aplicación sin estudios detallados previos. No se debería desplazar la intervención humana, dado que la precisión nunca es total, aunque se acerque bastante a ello. Por el momento, la experiencia práctica de más proyectos en esta línea de investigación, con la implicación de nosotros los archivistas, podrá revelar nuevas claves importantes.

Durante la conceptualización y el análisis de los distintos artículos que conformaron este trabajo, puedo decir que como archivista se abre un amplio panorama de oportunidad en la investigación, al delinear cómo podemos trabajar con los expertos en las ciencias computacionales. La revisión me permitió explorar múltiples metodologías, enfoques y abordajes. Más allá de esto me permite acercarme a esta en las diferentes formas que puedan adquirir sus implementaciones. Creo que en términos generales el visitar experiencias previas en términos de prototipos, pruebas de concepto y propuestas que da pie a la capacidad de formular proyectos de investigación con una serie de prácticas relacionadas.

En estos primeros inicios de la interacción, principalmente como archivistas nos hallamos en una fase de entendimiento y entrelazamiento en la implementación de estas tecnologías. Queda por ver estas propuestas o ideas del uso de la IA que tan precisas pueden ser, así como seguir elaborando desde la archivística qué más es posible proponer para sumar al desarrollo responsable de modelos computacionales. Por el momento nos queda articular estos nuevos conceptos y términos dentro de nuestro bagaje que permita formular nuevas ideas al respecto, como ya ha sucedido con otras tecnologías que se han incorporado en el pasado a nuestra profesión. En esta tesis se ha realizado una revisión de literatura a profundidad para reunir información que proporcione un marco de partida que estructure la relación entre la archivística y la inteligencia artificial. El presente marco de referencia presenta líneas de investigación en relación a las funciones y actividades posibles para la profesión archivística así como también su puesta en práctica en archivos. La protección de datos personales en el contexto de la IA requiere un análisis más profundo en casos de uso específico para poder generalizar su efectividad. Nuestra formación en la gestión de datos será crucial para establecer una participación más activa, pero debe entenderse que es el trabajo en conjunto con otras ciencias y disciplinas, el que podrá conducir a resultados cada vez más precisos, confiables y justos.

Referencias

ACUERDO del Consejo Nacional del Sistema Nacional de Transparencia, Acceso a la Información Pública y Protección de Datos Personales, por el que se aprueban los Lineamientos generales en materia de clasificación y desclasificación de la información, así como para la elaboración de versiones públicas. https://www.dof.gob.mx/nota_detalle_popup.php?codigo=5433280

Archivo General de la Nación. (2020). Guía de Anonimización de Datos Estructurados - Conceptos generales y propuesta metodológica. Recuperado 20 de julio de 2023, de https://www.archivogeneral.gov.co/sites/default/files/Estructura_Web/5_Consulte/Recursos/Publicacionees/2021-06-25_Guia_de_Anonimizacion.pdf

Álvarez Aguilar, C., Arcos, C., Figueroa Mora, K. M., Flores Garrido, M., García Velázquez, L. M., Gómez García, A., López Pineda, A., Mercado Ríos, J., Rodríguez Moran, M. F., & Tripp Gudiño, J. (2022). Estudio Longitudinal para el desarrollo de modelos predictivos de complicaciones de la diabetes melitus tipo 2. *Komputer Sapiens*, 14(3), 10-14. http://smia.mx/komputersapiens/descargar.php?file=ks143_3.88MB_compacta.pdf

Arnau Sabatés, L., & Sala Roca J. (2020). La revisión de la literatura científica: Pautas, procedimientos y criterios de calidad. Universitat Autònoma de Barcelona. Recuperado el 4 de Junio del 2022. Recuperado de: https://ddd.uab.cat/pub/recdoc/2020/222109/revliltcie_a2020.pdf. [Consultado el 1 de Septiembre del 2022]

Barranquero, J. (2022). INTRODUCCIÓN A LA ANONIMIZACION DE DATOS Técnicas y casos prácticos. Gobierno de España. https://datos.gob.es/sites/default/files/doc/file/introduccion_a_la_anonimizacion_de_datos-_tecnicas_y_casos_practicos_1.pdf

Canto Collí, L. (2017). Los archivos empresariales: un factor determinante en la economía del conocimiento. *Antrópica Revista De Ciencias Sociales Y Humanidades*, 3(5).. Recuperado de: <https://dialnet.unirioja.es/servlet/articulo?codigo=7423874>

Centro Nacional de Memoria Histórica & Archivo General de la Nación. (2022). Guía para la anonimización de datos e información no estructurada: Estándares y lineamientos técnicos. Recuperado 20 de julio de 2023, de <https://centrodememoriahistorica.gov.co/wp-content/uploads/2022/08/GUIA-DE-ANONIMIZACION.pdf>

Cerrillo Martínez, A., & Casadesús de Mingo, A. (2018). El impacto de la gestión documental en la transparencia de las Administraciones públicas: La transparencia por diseño. *Gestión y Análisis de Políticas Públicas*, 6-16. <https://doi.org/10.24965/gapp.v0i19.10515>

Cortés Alonso, V. (1982). Manual de Archivos Municipales.

Cruz Mundet, J., 2011. 1. Principios, términos y conceptos fundamentales. En: Administración de documentos y archivos. Textos fundamentales. [online] Coordinadora de Asociaciones de Archiveros, p.486. Recuperado de: <https://e-archivo.uc3m.es/bitstream/handle/10016/19730/archivos_cruz_2011.pdf?sequence=1&isAllowed=y> [Consultado el 1 de Septiembre del 2022]

Cámara de Diputados del H. Congreso de la Unión. Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental Disponible en: https://www.diputados.gob.mx/LeyesBiblio/abro/lftaipg/LFTAIPG_orig_11jun02.pdf

Cámara de Diputados del H. Congreso de la Unión (15 de junio del 2022). Ley General de Archivos. DOF 05-04-2022. Recuperado de: <https://www.diputados.gob.mx/LeyesBiblio/pdf/LGA.pdf>. [Consultado el 1 de Septiembre del 2022]

Cámara de Diputados del H. Congreso de la Unión (28 de enero de 2017). Ley General de Protección de Datos Personales en Posesión de Sujetos Obligados. DOF 26-01-2017. Recuperado de: <https://www.diputados.gob.mx/LeyesBiblio/pdf/LGPDPSO.pdf>.

Cámara de diputados del H. Congreso de la Unión. Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental Disponible en: <https://www.uv.mx/contraloria/files/2013/02/8.-Ley-Federal-de-Transparencia-y-Acceso-a-la-Informacion-Publica-Gubernamental.pdf>

Cámara de diputados del H. Congreso de la Unión.. Ley Federal de Protección de Datos Personales en Posesión de los Particulares. Disponible en: <https://www.diputados.gob.mx/LeyesBiblio/pdf/LFPDPPP.pdf>

Cámara de Diputados del H. Congreso de la Unión. Constitución Política de los Estados Unidos Mexicanos. Reforma del 6 de mayo del 2023. Disponible en: <https://www.diputados.gob.mx/LeyesBiblio/pdf/CPEUM.pdf>

Csányi, G. M., Nagy, D., Vági, R., Vadász, J. P., & Orosz, T. (2021). Challenges and open problems of legal document Anonymization. *Symmetry*, 13(8), 1490. <https://doi.org/10.3390/sym13081490>

Danks, D., & London, A. J. (2017). Algorithmic bias in autonomous systems. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*. <https://doi.org/10.24963/ijcai.2017/654>

del Campo Puerta, P. (2019). Archivo personal. Archivo personalizado. *Amoxthli*, 5, 59-73. <https://www.redalyc.org/journal/6157/615768192003/html/>

- Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems*, 14(3), 330-347. <https://doi.org/10.1145/230538.230561>
- Fuster Ruiz, F. (1999). Archivística, archivo. documento de archivo... Necesidad de clarificar los conceptos. *Anales de documentación*, 2, 103-120. Recuperado de: <https://revistas.um.es/analesdoc/article/view/2631>. [Consultado el 1 de Septiembre del 2022]
- Garat, D., & Wonsever, D. (2022). Automatic Curation of court documents: Anonymizing personal data. *Information*, 13(1), 27. <https://doi.org/10.3390/info13010027>
- Gardner, J., & Xiong, L. (2008). HIDE: An integrated system for health information DE-identification. 2008 21st IEEE International Symposium on Computer-Based Medical Systems. <https://doi.org/10.1109/cbms.2008.129>
- Gilli, J. J. (2017). La transparencia Como objetivo del desarrollo sostenible. The transparency as an objective the sustainable development. *Ciencias Administrativas*, (9), 006. <https://doi.org/10.24215/23143738e006>
- González Pedraza, J. A. (2010). Los archivos de empresas: un estudio comparativo. *Revista Andaluza de Archivos*, 2, 32. http://eprints.rclis.org/14203/1/archivos_de_empresas.pdf
- Hernández Sampieri, R. (2014). *Metodología de la Investigación* (6th ed., p. 365). McGraw-Hill / Interamericana Editores, S.A. DE C.V.
- Herrera, A. H. (1991). *Archivística general: Teoría y práctica* (5th ed.). Diputación Provincial de Sevilla.
- InterPARES 1 Project. s.f. InterPARES 1 Project. [online] Recuperado de: <http://www.interpares.org/ip1/ip1_documents.cfm?cat=atf> [Consultado el 1 de Agosto del 2022].
- InterPARES 2 Project. 2002. Project Summary. Recuperado de: http://web.archive.org/web/20220401032147/http://www.interpares.org/ip2/ip2_index.cfm. [Consultado el 1 de Septiembre del 2022]
- InterPARES 3 Project. 2007. Project Summary. Recuperado de: http://web.archive.org/web/20220419185335/http://www.interpares.org/ip3/ip3_index.cfm. [Consultado el 1 de Septiembre del 2022]
- InterPARES Trust. 2018. InterPARES Trust. Recuperado de: <http://interparestrust.org/>. [Consultado el 1 de Septiembre del 2022]
- Interpares Trust AI - Artificial Intelligence. 2021. Interparestrustai.Org. Recuperado de: <https://interparestrustai.org/>. [Consultado el 1 de Septiembre del 2022]

- Jo, E. S., & Gebru, T. (2020). Lessons from archives. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. <https://doi.org/10.1145/3351095.3372829>
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25.
- Lee, N. T., Resnick, P., & Barton, G. (2022). Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. Brookings. <https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/>
- Marciano, R., Lemieux, V., Hedges, M., Esteva, M., Underwood, W., Kurtz, M. & Conrad, M. (2018). Archival Records and Training in the Age of Big Data. In J. Percell , L. C. Sarin , P. T. Jaeger , J. C. Bertot (Eds.), *Re-Envisioning the MLS: Perspectives on the Future of Library and Information Science Education (Advances in Librarianship, Volume 44B, pp.179-199)*. Emerald Publishing Limited
- Martin, F., & Johnson, M. (2015). More Efficient Topic Modelling Through a Noun Only Approach. Proceedings of the Australasian Language Technology Association Workshop 2015, 111-115.
- Mena Mugica, Mayra M. 2020. "El Cambio De Paradigma En El Campo De La Archivística Desde Finales Del Siglo XX". En *El Proyecto Inter pares En América Latina Y El Caribe. Apuntes Sobre Archivos Digitales, Transparencia, Acceso A La Información Y Protección De Datos Personales*, 1er ed., 13-54. México: Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales. Recuperado de: <https://inai.janium.net/janium/Documentos/3801%20InterPARES.pdf>. [Consultado el 1 de Septiembre del 2022]
- Meystre, S. M., Friedlin, F. J., South, B. R., Shen, S., & Samore, M. H. (2010). Automatic de-identification of textual documents in the electronic health record: A review of recent research. *BMC Medical Research Methodology*, 10(1). <https://doi.org/10.1186/1471-2288-10-70>
- Mohammad, A. H., Alwada'n, T., & Al-Momani, O. (2016). Arabic text categorization using support vector machine, Naïve Bayes and neural network. *GSTF Journal on Computing (JoC)*, 5(1). <https://doi.org/10.7603/s40601-016-0016-9>
- Murphy, K. P. (2012). *Machine learning: A probabilistic perspective (Ser. Adaptive computation and machine learning)*. The MIT Press.

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association*, 18(5), 544–551. <https://doi.org/10.1136/amiajnl-2011-000464>

National Archives of Australia, The. (s.f.). Born-Digital File Format Standards. Recuperado de: <https://www.naa.gov.au/information-management/storing-and-preserving-information/preserving-information/born-digital-file-format-standards>. [Consultado el 1 de septiembre del, 2022]

Peters, M. D., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, 13(3), 141-146. Recuperado de: <https://doi.org/10.1097/xeb.0000000000000050>. [Consultado el 1 de Septiembre del 2022]

Rangel, E. L., Ramirez Roa, D., & Rodriguez, C. (2020). GUÍA DE ANONIMIZACIÓN DE DATOS ESTRUCTURADOS: Conceptos generales y propuesta metodológica. ARCHIVO GENERAL DE LA NACIÓN. https://www.archivogeneral.gov.co/sites/default/files/Estructura_Web/5_Consulte/Recursos/Publicacionees/Guia_de_Anonimizacion-min.pdf

Real Academia Española. (s.f.). Inteligencia Artificial. En: Diccionario de la lengua española. Recuperado de <https://dle.rae.es/inteligencia#2DxmhCT>. [Consultado el 1 de septiembre de 2022]

Real Academia Española. (s.f.). Privacidad. En: Diccionario panhispánico del español jurídico. Recuperado de <https://dle.rae.es/inteligencia#2DxmhCT>. [Consultado el 20 de abril de 2023]

Reshamwala, A., Mishra, D., & Pawar, P. (2013). REVIEW ON NATURAL LANGUAGE PROCESSING. – *Engineering Science and Technology: An International Journal (ESTIJ)*, 3(1), 1. https://www.researchgate.net/profile/Alpa-Reshamwala/publication/235788362_REVIEW_ON_NATURAL_LANGUAGE_PROCESSING/links/00463516276f412048000000/REVIEW-ON-NATURAL-LANGUAGE-PROCESSING.pdf

Ríos, P. (2008). La importancia de la organización y conservación de un archivo personal. *Revista Universidad de Sonora*, 23, 52.

Russell, S., & Norvig, P. (2009). *Artificial intelligence: A modern approach* (3rd ed.). Pearson.

Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Networks*, 61, 85–117. Recuperado de: <https://doi.org/https://www.sciencedirect.com/science/article/abs/pii/S0893608014002135?via%3Dihub>. [Consultado el 1 de Septiembre del 2022]

Skansi, S. (2018). En. Introduction to deep learning: From logical calculus to artificial intelligence (Ser. Undergraduate Topics in Computer Science, pp. 1–187). Springer.

The National Archives. (s.f.). Born-digital records and metadata. Recuperado de: <https://www.nationalarchives.gov.uk/information-management/manage-information/digital-records-transfer/what-are-born-digital-records/>. [Consultado el 1 de Septiembre del 2022]

Weka 3 - Data mining with open source machine learning software in Java. (s.f.). Department of Computer Science: University of Waikato. <https://www.cs.waikato.ac.nz/ml/weka/>

Voutssas M., J., & Barnard Amozorrutia, A. (2014). Glosario de preservación archivística digital versión 4.0 [Ebook]. UNAM, Instituto de Investigaciones Bibliotecológicas y de la Información. Recuperado de https://iibi.unam.mx/archivistica/glosario_preservacion_archivistica_digital_v4.0.pdf. [Consultado el 1 de Septiembre del 2022]

ANEXO

Ejes Temáticos		
Relevancia	1. Archivos, archivística, gestión documental e IA	2. Regulación, y tecnología
Considerable	<p>Andresen, H. (2019). A discussion frame for explaining records that are based on algorithmic output. <i>Records Management Journal</i>, 30(2), 129-141. https://doi.org/10.1108/rmj-04-2019-0019</p> <p>Bunn, J. (2020). Working in contexts for which transparency is important. <i>Records Management Journal</i>, 30(2), 143-153. https://doi.org/10.1108/rmj-08-2019-0038</p> <p>Büttner, G. (2019). Auto-classification in an international organization: Report from a feasibility study. <i>Comma</i>, 2017(2), 15-26. https://doi.org/10.3828/comma.2017.2</p> <p>Chabin, M. (2020). The potential for collaboration between AI and archival science in processing data from the French great national debate. <i>Records Management Journal</i>, 30(2), 241-252.</p> <p>Colavizza, G., Blanke, T., Jeurgens, C., & Noordegraaf, J. (2021). Archives and AI: An overview of current debates and future perspectives. <i>Journal on Computing and Cultural Heritage</i>, 15(1), 1-15.</p> <p>Makhlouf Shabou, B., Tièche, J., Knafou, J., & Gaudinat, A. (2020). Algorithmic methods to explore the automation of the appraisal of structured and unstructured digital data. <i>Records Management Journal</i>, 30(2), 175-200. https://doi.org/10.1108/rmj-09-2019-0049</p> <p>Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsoupova, T., & Stuart, K. (2018). More</p>	<p>Hutchinson, T. (2018). Protecting privacy in the archives: Supervised machine learning and born-digital records. 2018 IEEE International Conference on Big Data (Big Data).</p> <p>Hamon, R., Junklewitz, H., Sanchez, I., Malgieri, G., & De Hert, P. (2022). Bridging the gap between AI and Explainability in the GDPR: Towards trustworthiness-by-Design in automated decision-making. <i>IEEE Computational Intelligence Magazine</i>, 17(1), 72-85Trustworthiness-by-Design in Automated Decision-Making</p> <p>Jaillant, L., & Caputo, A. (2022). Unlocking digital archives: Cross-disciplinary perspectives on AI and born-digital data. <i>AI & SOCIETY</i>, 37(3), 823-835. https://doi.org/10.1007/s00146-021-01367-x</p> <p>Jo, E. S., & Gebru, T. (2020). Lessons from archives. <i>Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency</i>. https://doi.org/10.1145/3351095.3372829</p>

human than human? Artificial intelligence in the archive. *Archives and Manuscripts*, 47(2), 179-203. <https://doi.org/10.1080/01576895.2018.1502088>

Vellino, André, and Inge Alberts. 2016. 'Assisting the Appraisal of E-Mail Records with Automatic

Classification'. *Records Management Journal* 26 (3): 293–313