

Visualizar la desaparición forzada: grafos de conocimiento para la búsqueda y la verdad

Visualizing Forced Disappearance: Knowledge Graphs for Search and Truth

Mariana Esther Martínez Sánchez,*
Tamara Donají Valencia López**



Dossier

Transformación digital e inteligencia artificial:
Debates actuales, dilemas futuros desde
las ciencias sociales

Licencia Creative Commons Atribución-NoComercial
(CC BY-NC) 4.0 Internacional

Perfiles Latinoamericanos, 34(67) | 2026 | e-ISSN: 2309-4982

DOI: dx.doi.org/10.18504/pl3467-001-2026

Recibido: 5 de agosto de 2025

Aceptado: 4 de diciembre de 2025

Resumen

El sistema Angelus fue desarrollado para integrar, analizar y visualizar información sobre desapariciones forzadas durante la “guerra sucia” (1964-1985) en México. Utilizando grafos de conocimiento y una ontología computacional, permite estructurar datos dispersos y contradictorios provenientes de archivos oficiales, testimonios y otras fuentes. Su diseño responde a criterios de audibilidad, trazabilidad y centralidad del juicio humano, priorizando la utilidad para víctimas, familiares e instituciones de búsqueda. Este artículo presenta los retos técnicos, metodológicos y éticos enfrentados durante su desarrollo, así como las lecciones aprendidas para futuros proyectos de tecnología social aplicados a violaciones graves de derechos humanos.

Palabras clave: desaparición forzada, guerra sucia, derechos humanos, grafos de conocimiento, tecnología social, archivos históricos, verdad y justicia, anotación documental, búsqueda de personas.

Abstract

The Angelus system was developed to integrate, analyze, and visualize information on forced disappearances during Mexico’s “dirty war” (1964-1985). Using knowledge graphs and computational ontology, it structures fragmented and contradictory data from official archives, testimonies, and other sources. Its design prioritizes auditability, traceability, and human judgment, emphasizing usefulness for victims, families, and search institutions. This article presents the technical, methodological, and ethical challenges faced during its development,

* Doctora en Ciencias Biomédicas por la Universidad Nacional Autónoma de México | Investigadora en Ciencias Médicas en el Instituto Nacional de Enfermedades Respiratorias “Ismael Cosío Villegas” | maresther.martinez@iibiomedicas.unam.mx | <https://orcid.org/0000-0002-6115-1449>

** Licenciada en Estudios Latinoamericanos por la Universidad Nacional Autónoma de México | Investigador independiente | valencia.donaji.95@gmail.com | <https://orcid.org/0009-0004-3168-1416>

as well as lessons learned for future social technology projects aimed at addressing serious human rights violations.

Keywords: forced disappearance, dirty war, human rights, knowledge graphs, social technology, historical archives, truth and justice, document annotation, missing persons search.

Introducción

Durante el periodo conocido como “guerra sucia” (1964-1985), el Estado mexicano llevó a cabo una represión sistemática y planificada contra grupos armados y civiles opositores al régimen. Entre las técnicas utilizadas destacan la detención ilegal y la desaparición forzada de integrantes de organizaciones políticas, así como de sus familiares, acciones conocidas como circuito de detención-desaparición (Aguayo, 2001; Vicente Ovalle, 2019). El Mecanismo para la Verdad y el Esclarecimiento Histórico (MEH) identificó 1166 víctimas de desaparición forzada: 649 transitorias y 517 que siguen desaparecidas (MEH, 2024a). La incertidumbre sobre su destino, sumada a la falta de justicia y de acceso a la verdad, genera un daño persistente tanto a sus familiares como a la sociedad en su conjunto.

Entre las instituciones gubernamentales que participaron como represoras en estos circuitos se encuentran la Secretaría de la Defensa Nacional (Sedena) y diversas dependencias de la Secretaría de Gobernación (Segob), como la Dirección General de Investigaciones Políticas y Sociales (DGIPS) y la Dirección Federal de Seguridad (DFS) (Calderón & Cedillo, 2012; MEH, 2024b). Dado que estas graves violaciones a los derechos humanos fueron perpetradas por instituciones del Estado con el respaldo de otras entidades gubernamentales, existe una vasta documentación oficial de sus acciones, hoy dispersa en los archivos de distintas dependencias (Cedillo & Valenzuela, 2014).

Para esclarecer las desapariciones ocurridas durante la “guerra sucia”, no basta con abordar los casos de manera individual. Es imprescindible integrar la información disponible de manera sistemática, a fin de identificar patrones y detectar vacíos críticos de conocimiento (Caridi *et al.* & Somigliana, 2011; Carrington *et al.*, 2005). Para satisfacer esta necesidad, la Comisión Nacional de Búsqueda (CNB), en colaboración con el Consejo Nacional de Humanidades, Ciencias y Tecnologías (Conahcyt), desarrollaron el sistema y grafo de conocimiento¹ An-

¹ Un grafo de conocimiento o red semántica, representa una red de entidades del mundo real (personas, objetos, eventos, conceptos, por ejemplo) y sus relaciones.

gelus, cuyo objetivo ha sido sistematizar y analizar la información contenida en archivos históricos relacionados con violaciones graves a los derechos humanos, particularmente en relación con la desaparición forzada durante la “guerra sucia” (Conahcyt, 2024a).

Una característica esencial de la desaparición de personas es la carencia de información: por definición, se desconoce el paradero de la víctima. Sin embargo, para que alguien sea considerado desaparecido, es necesario que otra persona lo esté buscando, lo que implica una conciencia explícita —a nivel individual e institucional— de dicha falta de información (Puerto *et al.*, 2021). Por ejemplo, debido a la naturaleza clandestina de la guerrilla urbana, muchos de sus integrantes cambiaron de identidad y se desvincularon de sus familias, lo que implica que algunas personas podrían haber ingresado al circuito de detención-desaparición sin que siquiera se tenga noticia de su existencia (Rodríguez Kuri, 2024). Esto genera problemáticas vinculadas a lo que sabemos, a lo que sabemos que no sabemos y, de forma más compleja, a lo que no sabemos que no sabemos. Sabemos que ignoramos la ubicación de los desaparecidos, o que existen acervos de los cuales desconocemos su contenido, pero la mayor dificultad radica en identificar aquello que ignoramos por completo.

El uso de sistemas de información y herramientas tecnológicas en el ámbito de los derechos humanos se ha vuelto central para documentar violaciones, analizar patrones de violencia y sustentar procesos de verdad y justicia, pasando de los proyectos de documentación utilizando bases de datos y el desarrollo de estándares computacionales (Levin & Hamber, 1997; ANM, 2017; Dueck *et al.*, 2001), a plataformas complejas que integran análisis estadístico avanzado, aprendizaje automático, información espacial y grafos de conocimiento (Price *et al.*, 2009; Guberek & Hedstrom, 2017; Green & Ball, 2019; Koho *et al.*, 2020; HURIDOCs, 2025; Ushahidi, s. f.). Sin embargo, estas tecnologías también han sido instrumentalizadas para su violación, como lo evidencian el uso de tarjetas perforadas en el Holocausto y las formas contemporáneas de autoritarismo algorítmico basadas en vigilancia masiva y censura automatizada. Por ello, es fundamental que las organizaciones sociales y las comunidades afectadas se apropien críticamente de estas herramientas y que las decisiones permanezcan en manos de personas.

El sistema Angelus facilita la integración, homologación y explotación de datos dispersos en múltiples fondos documentales. El flujo de trabajo incluye la digitalización, selección y anotación de documentos para generar un grafo de conocimiento, el cual permite la visualización y análisis de la información contenida en él para identificar patrones de represión, reconstruir trayectorias de personas desaparecidas y ubicar redes institucionales (Angelus, 2022a, 2022b). El sistema Angelus no es solo una herramienta de investigación para



instituciones y especialistas, representa un esfuerzo por hacer accesible el conocimiento archivístico a los familiares de personas desaparecidas y a la sociedad en general, contribuyendo así al derecho a la verdad. Su desarrollo responde a la necesidad de abordar el fenómeno de la desaparición no solo desde los casos individuales, sino desde una perspectiva estructural que permita comprender la magnitud y lógica de la represión estatal. El sistema Angelus pone en el centro de su funcionamiento a las personas, reconociendo que la responsabilidad de las decisiones recae en los humanos; reconociendo los derechos de los familiares a conocer la verdad y a participar en los procesos de esclarecimiento y obtener justicia.

Ley de amnistía

En el caso de la “guerra sucia”, existen diversos tipos de documentos: archivos del circuito de detención-desaparición emitidos por las autoridades responsables de las acciones de contrainsurgencia, tales como la Sedena, DFS, DGIPS, etc. (Archivo General de la Nación, 2020; Artículo 19, 2015); fondos de instituciones que formaron parte de los circuitos de detención durante el periodo, como hospitales, servicios médico-forenses (SEMEFO), cárceles, entre otros (Cedillo & Valenzuela, 2014); archivos generados por instituciones que han investigado los casos, como la Comisión Nacional de Derechos Humanos (CNDH) (CNDH, 2001), la Fiscalía Especial para Movimientos Sociales y Políticos del Pasado (FEMOSPP) (FEMOSPP, 2006) y la Comisión de la Verdad del Estado de Guerrero (ComVerdad) (ComVerdad, 2014); múltiples archivos de colectivos y personales de familiares y sobrevivientes (Artículo 19, 2021); además de las fuentes hemerográficas de la época y los testimonios recabados por la propia CNB.

No obstante, resulta relevante mencionar que tener el archivo no significa tenerlo todo o tener “la verdad”. Es siempre indispensable tomar en cuenta el contexto propio de los archivos, reconocer a quienes emiten la información, a quién está dirigida, dónde fue emitida, si los documentos contienen fecha, e incluso los formatos. Un ejemplo es el de un informe presentado en 1979 por Óscar Flores Sánchez, el entonces procurador general de la república, basado en el expediente de la DFS sobre “el seguimiento de 314 militantes a nivel nacional” (DFS, 1971), en donde la mayoría de la información es falsa y muy posiblemente generada después de los hechos que supuestamente reporta (Gamiño Muñoz, 2020).

El contexto en el que este informe se hace público resalta dos particularidades. Por un lado, Rosario Ibarra estaba impulsando el proceso penal por la desapa-

rición de su hijo Jesús Piedra Ibarra, detenido-desaparecido desde el 18 de abril de 1975 en Monterrey, Nuevo León (MEH, 2024a). La visibilidad política que ya tenía Rosario Ibarra para 1979 era de nivel internacional, sin dejar de lado el apoyo de otros familiares con quienes había conformado el Comité Nacional Pro Defensa de Presos, Perseguidos, Desaparecidos y Exiliados Políticos de México, conocido actualmente como Comité ¡Eureka! Por otro lado, el 28 de septiembre de 1978 se decreta la Ley de Amnistía por el entonces presidente de la república, José López Portillo (1978), cuyo primer artículo señala que:

Se decreta amnistía en favor de todas aquellas personas en contra de quienes se haya ejercitado acción penal, ante los Tribunales de la Federación o ante los Tribunales del Distrito Federal en materia de fuero común, hasta la fecha de entrada en vigor de la presente Ley, por los delitos de sedición, o porque hayan invitado, instigado o incitado a la rebelión, o por conspiración u otros delitos cometidos formando parte de grupos e impulsados por móviles políticos con el propósito de alterar la vida institucional del país, que no sean contra la vida, la integridad corporal, terrorismo o secuestro.

No obstante, la aplicación de esta ley implicaba que los cientos de personas desaparecidas por motivos políticos a manos de autoridades locales y federales debían encontrarse detenidas en centros penitenciarios, lo cual no era el caso. En 1992, la CNDH reafirmó su postura respecto al informe presentado por Óscar Flores Sánchez (Ballinas, 2001), señalando que no existía evidencia que respaldara los hechos allí descritos. En algunos casos, el informe indicaba fechas de fallecimiento que resultaban insostenibles, ya que las personas seguían con vida en esos momentos, e incluso existían otros documentos oficiales que contradecían dichas versiones. En otros casos, se afirmaba que las personas se encontraban prófugas o que habían sido ejecutadas por sus propios compañeros de militancia, todo ello sin fundamento documental verificable, más allá de supuestos informes emitidos por la DFS, sin fecha de emisión ni firma al calce.

A continuación se presenta un ejemplo de estos documentos. Adolfo Tecla Parra tenía 15 años cuando fue detenido-desaparecido el 3 de junio de 1975, junto con su hermana Violeta. Ella fue posteriormente presentada ante las autoridades y trasladada al reclusorio femenino de Santa Martha Acatitla, donde permaneció detenida por más de dos años antes de ser liberada en virtud de la Ley de Amnistía. De Adolfo, en cambio, no se volvió a tener noticia. En el expediente de la DFS con el seguimiento de 314 militantes a nivel nacional presentado en 1979, se reporta que Adolfo Tecla Parra, de 15 años, fue detenido el 28 de agosto de 1975 en una casa de seguridad en Ciudad Nezahualcóyotl, pero que logró escapar y se encontraba prófugo.



Figura 1A. Parte del informe presentado por Óscar Flores Sánchez en 1979

287.- TECLA PARRA ADOLFO

Miembro de la llamada Liga Comunista "23 de Septiembre".

El 28 de agosto de 1975, fue localizado en una "casa de seguridad" que tenía la Liga en Ciudad Netzahualcóyotl, Méx., junto con otros de sus compañeros, encontrándoseles en su poder armas de grueso calibre y bombas de fabricación casera, así como una imprenta. Al tratar de --detenerlos, abrieron fuego contra las fuerzas públicas, logrando herir a dos agentes policiacos, a los cuales se tuvo que trasladar a un hospital, dejando a uno para vigilar a los detenidos.

Poco después, se encontró a este elemento herido y amordazado, manifestando que los detenidos habían logrado huir, llevando consigo -- algunas armas que tenían escondidas en el piso. Desde esa fecha, -- se ignora el paradero de ADOLFO TECLA PARRA, el cual es buscado -- por las distintas corporaciones policiacas para que responda por la -- comisión de varios delitos.

Figura 1B. Informe de la DFS con el que Óscar Flores Sánchez justifica lo que señaló en el informe de 1979

TECLA PARRA ADOLFO ²⁸⁹⁻
345 20001

Este individuo se encuentra prófugo de la justicia y se sabe que es miembro de la Liga Comunista "23 de Septiembre", a nombre de cuyo grupo ha realizado múltiples hechos delictivos.

El 28 de agosto de 1975, fue localizado en una casa de seguridad en Ciudad Netzahualcóyotl, Estado de México, en unión de otros miembros de dicho grupo subversivo encontrándoseles en su poder armas de grueso calibre, propaganda y una imprenta en la que imprimían el órgano de difusión de la Liga Comunista "23 de Septiembre", así como innumerables -- artefactos explosivos de fabricación casera.

Al intentarse detenerlos, agredieron a los agentes de la autoridad, lesionando a dos de ellos, por lo que fue -- preciso trasladarlos a un hospital habiendo dejado a uno de ellos vigilando a Adolfo Tecla Parra y a otro miembro del mismo grupo cuyo nombre no fue revelado.

Al regresar con refuerzos los agentes del orden -- público 30 minutos después a dicha casa, encontraron muerto al Agente que quedó encargado de custodiarlos en el interior del baño, por lo que se supone que fue atacado por este sujeto y su compañero huyendo y llevándose consigo varias armas que por un orificio encontrado en el suelo, se supone escondían ahí.

A partir de esa fecha, diversas corporaciones policiacas han tratado de lograr su captura sin que hasta la -- fecha se haya podido localizar.



ARCHIVO GENERAL
DE LA NACIÓN

Puede observarse que, aunque reproduce el formato general de los informes emitidos por la DFS, carece de fecha y de firma al calce (figuras 1A, 1B). Sin embargo, un documento interno de la misma DFS emitido el 3 de junio de 1975 reporta la detención de Adolfo, junto con su hermana Violeta, en la colonia Clavería (figura 1C). Este caso ilustra cómo las autoridades manipularon la información con el fin de evadir su responsabilidad en graves violaciones a derechos humanos, incluso en casos que involucraron a menores de edad.

Figura 1C. Parte del informe de la DFS emitido el 3 de junio de 1975 en el que reportan la detención de Adolfo Tecla Parra

11-235
L-291
200001

D. F. S. -3-VI-75.

LIGA COMUNISTA "23 DE SEPTIEMBRE"

"BRIGADA ROJA"

El día de hoy elementos de esta Dirección Federal de Seguridad y de la División de Investigaciones para la Prevención de la Delincuencia, continuaron el interrogatorio de ALFONSO TECLA PARRA (a) "Rafael", quien entre otros asuntos señaló un punto de contacto frente al jardín de Clavería a inmediaciones de la parte posterior de la Preparatoria Popular de Teacubán, donde aproximadamente a las 18,00 horas fueron aprehendidos VIOLETA TECLA PARRA (a) "Elena" o "Laura" de 19 años de edad, ADOLFO TECLA PARRA (a) "Jesús" de 14 años, ambos pertenecientes a la "Brigada Roja" de la Liga Comunista "23 de Septiembre"; así como a JUSTINO SERRATO HERRERA quien trabaja en la Empresa "Tesoros" como troquelador, ubicada en la calle Pirámide en Naucalpan, Méx., y a su hermano VICENTE de 18 años de edad, quien es estudiante en el Colegio de Bachilleres; estos últimos con domicilio en la calle de Pimas No. 34, Col. Las Trancas en Acapatzalco, D.F.

Como se puede notar de esta serie de documentos, es necesario un sistema de manejo de información capaz de gestionar información archivística compleja y contradictoria, propia de contextos de represión estatal. Esto implica los siguientes requerimientos:

- El sistema debe ser auditable, escalable y diseñado bajo principios de seguridad, confidencialidad y respeto a los derechos humanos.



- Gestión de contradicciones y mentiras: El sistema debe permitir registrar versiones divergentes sobre un mismo hecho, sin suprimir las contradicciones, para facilitar su análisis crítico.
- Manejo de datos inciertos o incompletos: Debe ser robusto ante diversos niveles de certeza sobre fechas, lugares, nombres y demás datos clave, incluyendo rangos, estimaciones o variantes posibles.
- Soporte documental obligatorio: Todo enunciado registrado en el sistema debe estar vinculado a una fuente documental identificable, con metadatos suficientes para su trazabilidad.
- Capacidad de inferencia: El sistema debe poder relacionar información dispersa para generar trayectorias de personas, vínculos institucionales o patrones de actuación.
- Visualización y comunicación accesible: La información procesada debe poder presentarse de manera comprensible y útil para familiares de personas desaparecidas, permitiendo su consulta, descarga o integración en informes.
- Centralidad del juicio humano: El sistema no sustituye las decisiones humanas, dejando en todo momento el juicio final en manos de las personas usuarias.

Datos y lecciones aprendidas

La primera versión del sistema Angelus consistía en una hoja de cálculo elaborada por Javier Yankelevich, quien entonces fungía como Director de Operaciones de Búsqueda en la CNB. Este archivo contenía columnas con parámetros que indicaban “nombre”, “número de expediente”, “víctima”, “testigo” o “perpetrador”; sin embargo, existían múltiples limitaciones para trabajar de manera colaborativa, acceder a los documentos fuente o establecer relaciones (Santiago, 2022). Para resolver estas dificultades, se sugirió usar un enfoque basado en grafos de conocimiento, generando un sistema que permitiera resolver los requerimientos.

En 2019 inició la colaboración entre la CNB y el Conahcyt. De 2020 a 2024 se llevó a cabo el desarrollo por parte de un grupo de trabajo que incluyó personal e investigadores de la CNB, el Conahcyt, el Centro de Investigación en Matemáticas (CIMAT), el Centro de Investigación y Docencia Económicas (CIDE), el Instituto Nacional de Enfermedades Respiratorias (INER) y la Universidad Autónoma de México (UNAM) (Conahcyt, 2024a). De 2022 a 2023, el Conahcyt financió el proyecto en el CIDE. En junio de 2024, el sistema Angelus, incluyendo datos, manuales y código, fue entregado por el Conahcyt a la CNB, la cual se responsabiliza de su operación desde esa fecha (Conahcyt, 2024b).

Desarrollo

El desarrollo del proyecto implicó el diseño e implementación de un flujo de trabajo y las piezas de *software* correspondientes (figura 2). En general, el flujo de trabajo consiste en la digitalización de los archivos para convertirlos en objetos digitales;² estos últimos son organizados en un repositorio de documentos, a partir del cual son anotados de manera manual extrayendo las entidades y relaciones mencionados entre ellos. A partir de esto se visualizan los datos tanto en forma de grafo como en reportes. Para lograr esto se utilizaron varias piezas de *software*:

- *Software* de digitalización.
- Repositorio de documentos: Una modificación de Dataverse (The Dataverse Project, 2025) junto con un gestor de identificadores.³ Estos se encuentran envueltos en una API REST⁴ desarrollada por Eduardo Herrera que permite la generación de nodos del grafo correspondientes a documentos.
- Repositorio de grafo de conocimiento: Una instancia de Apache Jena Fuseki (Apache Jena Project, 2025), una base de datos de grafos. Dicha instancia se encuentra envuelta por una aplicación que permite su consulta por medio de métodos web de acuerdo a los permisos y la identidad de los usuarios.
- Anotador: Desarrollado específicamente para el proyecto por Víctor Mireles utilizando la biblioteca de desarrollo de interfaces Vue.js. Este sistema muestra un formulario cuyos campos y opciones están determinados por la ontología⁵ y el contenido actual del grafo; permite agregarle nuevos nodos y aristas.
- Visualizador de redes y reportes: Desarrollado específicamente para el proyecto por Sebastián Casillas, utilizando también Vue.js. El visualizador permite a la persona usuaria navegar el grafo a partir de entidades de su

² Un *objeto digital* es cualquier unidad de información que existe en formato electrónico y puede ser almacenada, procesada o transmitida por sistemas digitales. Incluye tanto el contenido (texto, imagen, audio, video, datos) como los metadatos (título, autor, fecha de creación, etc.) que lo describen y permiten su identificación y gestión.

³ Un *repositorio de documentos* es un sistema donde los archivos se almacenan de forma centralizada y cada documento recibe un identificador único que permite organizarlo, rastrearlo y recuperarlo sin ambigüedades.

⁴ Una API REST es una interfaz que permite que dos sistemas se comuniquen a través de internet usando reglas simples basadas en el protocolo HTTP.

⁵ Una *ontología computacional* es una representación formal y estructurada del conocimiento sobre un dominio específico que define los conceptos, relaciones entre ellos y las propiedades y reglas que describen cómo se comportan para que sean manejables en un sistema computacional.



interés, y generar reportes al momento sobre toda la información contenida sobre una entidad en particular en el grafo.

En todos los casos se implementaron controles de acceso⁶ basados en OAuth 2.0 para garantizar la seguridad de los datos. Además, con el apoyo de Víctor Muñiz y Oscar de la Cruz Echeveste del CIMAT y Juan Carlos González Aguilar de la UNAM se generaron varios flujos de datos de Reconocimiento Óptico de Caracteres (OCR) para convertir texto contenido en imágenes o documentos escaneados en texto digital editable y procesable por computadora, y procesamiento de lenguaje natural (NLP, por sus siglas en inglés) para extracción de entidades (De la Cruz Echeveste, 2023; González Aguilar, 2022). Como parte del compromiso del proyecto con el acceso a la ciencia, los datos y código se encuentran disponibles públicamente (Conahcyt, 2024a).

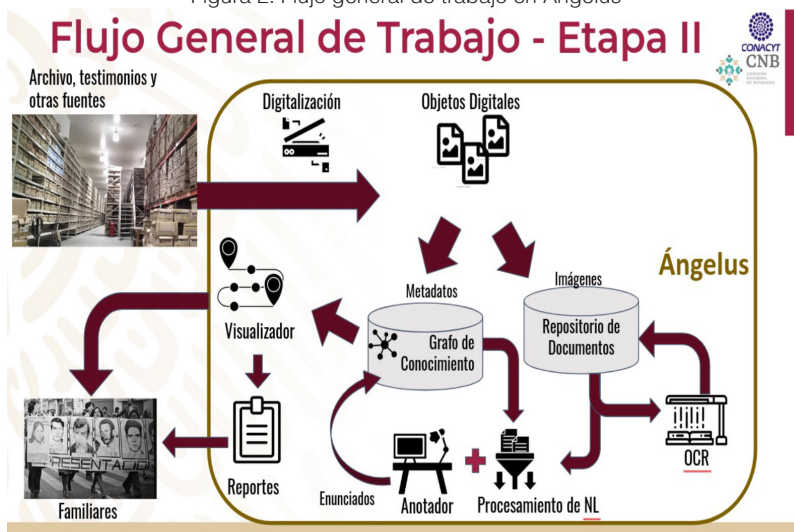
Durante el uso del sistema, se encontraron diversas dificultades: los mecanismos de corrección son limitados. El desarrollar *software* es un proceso altamente complejo, por lo que es recomendable reutilizar y adecuar herramientas ya existentes. El ligado de entidades⁷ es un problema no trivial, ya que es necesario saber cuándo distintos documentos se refieren a la misma entidad. Esto se ve dificultado por las características de los datos de origen. Para solucionar parcialmente esto, se buscaron anotadores expertos en el tema y se corrieron múltiples procesos de deduplicación de entidades. Las características de los documentos originales (edad, tipografía, daños, etc.) dificultan el uso de OCR y NLP.

Durante el desarrollo se puso en el centro a las personas. Se priorizó la utilidad práctica para los usuarios finales, desde analistas hasta familiares, dentro de las limitaciones de recursos disponibles. Se hizo especial énfasis en que la información contenida era sobre personas víctimas de violaciones a derechos humanos, y por lo tanto debe de ser tratada con sensibilidad y empatía. Además, por diseño se considera que las decisiones respecto a los casos, incluyendo la anotación, el análisis y la implementación de acciones de búsqueda, deben ser tomadas por las personas, de tal manera que el sistema sea una ayuda en los procesos. En ese aspecto, queremos resaltar que el sistema Angelus NO es una máquina de la verdad.

⁶ Los *controles de acceso* son mecanismos que regulan quién, y en qué condiciones, puede usar un sistema.

⁷ El *ligado de entidades (entity linking)* es el proceso de identificar menciones de entidades (personas, objetos, eventos, conceptos, por ejemplo) dentro de un texto y vincularlas con una entidad única en la base de conocimiento.

Figura 2. Flujo general de trabajo en Angelus



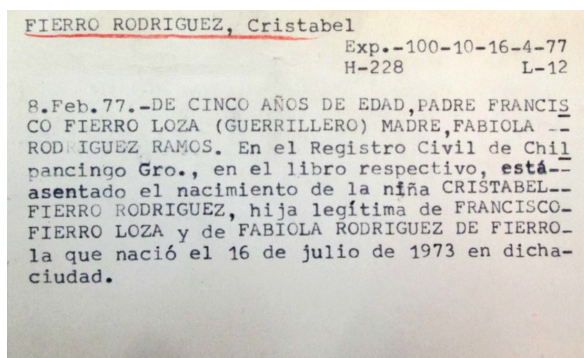
Ontología

Para poder generar el grafo de conocimiento del Angelus fue necesario diseñar una ontología computacional. Una ontología computacional es una especificación formal y legible por máquina de los conceptos, relaciones y reglas lógicas que definen la estructura y el significado de los datos en un dominio (Lanthaler *et al.*, 2014; Schneider *et al.*, 2009), en este caso, el circuito detención-desaparición en la “guerra sucia”. La ontología se compone de clases (o conceptos), que son categorías abstractas de cosas (p. ej., Persona, Lugar), y puede incluir subclases como Familiar y Víctima. Las propiedades describen cómo se relacionan las instancias entre sí o con datos literales. Las propiedades de objeto definen relaciones entre entidades (p. ej., estuvoEn), mientras que las propiedades de datos vinculan entidades a valores (p. ej., tieneNombrePropio). Los axiomas y las reglas son restricciones lógicas que establecen condiciones semánticas, como “Toda Víctima es una Persona”. Las instancias son entidades específicas que pertenecen a estas clases y subclases, a las cuales se les asignan identificadores. Por ejemplo, Adolfo Tecla Parra y Violeta Tecla Parra son instancias de la clase Persona, subclase Víctima, con propiedad de datos tieneNombrePropio: Adolfo y tieneNombrePropio: Violeta, respectivamente. Entre los dos existe una propiedad de objeto que describe la relación “Adolfo Tecla Parra esHermanoDe Violeta Tecla Parra”. Para saber más sobre la ontología del sistema Angelus se recomienda revisar el artículo de Mireles Chávez *et al.* (2021).



Una ontología computacional, si bien es potente para estructurar y razonar sobre datos, presenta varias limitaciones. Puede tener dificultades para representar conocimiento cambiante o ambiguo, verse afectada por las diferentes conceptualizaciones, vocabulario y experiencia de quien la diseña y usa y, al depender de la lógica formal, es por naturaleza rígida y no siempre capta los matices de las relaciones, emociones y el contexto (Blaser, 2013; D'Ignazio & Klein, 2023). Por ejemplo, la ficha de la DFS de Cristabel Fierro Rodríguez establece que tiene cinco años y es hija legítima de Francisco Fierro Loza (figura 3); sin embargo, esta anotación, aunque es fácticamente cierta, no incluye el contexto de violencia que existía en Guerrero ni el malestar que genera que una niña de cinco años fuera vigilada por la DFS. Otro ejemplo fue la decisión de usar sexo en lugar de género como propiedad, ya que ese es el campo que se usa en los documentos de la época, siendo conscientes de que esta decisión implicó ignorar la diferencia entre sexo biológico y el género, lo cual no es trivial dada la victimización de las disidencias sexo-genéricas en el periodo (MEH, 2024a).

Figura 3. Ficha de la DFS sobre Cristabel Fierro Rodríguez de cinco años de edad. (08/feb/77)



Digitalización

En paralelo al diseño del sistema, se buscó acceder a las diversas fuentes documentales para generar un inventario, organizar los documentos y digitalizarlos. Una de las principales dificultades fue negociar, concretar y mantener el acceso a los acervos documentales para la digitalización. En más de un caso se negoció el acceso a los documentos, pero no fue posible la digitalización o se retiró el acceso al acervo. Si bien no se tuvo acceso a la totalidad de los acervos, se digitalizó más del 80% del archivo histórico de la FEMOSPP; se recibió

una copia del archivo de ComVerdad; parte del archivo de Amnistía Internacional; archivo de un panteón de Guadalajara, otro de Sinaloa y del Panteón Civil Dolores en Ciudad de México, etc. (Angelus, 2022a).

Otra dificultad fue el tamaño de los acervos documentales y la falta de instrumento de consulta o estándares para nombrar y organizar los documentos, que permitieran conocer el contenido de los archivos resguardados. Esto implicó que para muchos fondos documentales fue necesario hacer un inventario previo a la digitalización. A esto se sumaron las dificultades físicas de la digitalización, como el acceso a escáneres de buena calidad, la presencia de grapas, dobladuras, etc., y el daño que volvió a los documentos sumamente frágiles.

Como parte del proceso de digitalización se buscó transformar imágenes de documentos digitalizados en texto editable y procesable por computadora utilizando herramientas de OCR y NLP, las cuales fueron desarrolladas en colaboración con el CIMAT y la UNAM. Este proceso fue altamente complejo debido a las características de los documentos originales: elaborados en máquina de escribir; con anotaciones a mano y rayaduras; han sido dañados a lo largo de los años; la rotación durante la digitalización es variable, y hay errores de origen. Los resultados fueron modestos, aunque se logró la digitalización a texto de los ficheros de la DFS y las “hojas amarillas” del SEMEFO-DF (De la Cruz Echeveste, 2023; González Aguilar, 2022).

Una consideración necesaria durante la digitalización, tanto en la captura de imágenes como en la transformación de estas a texto, fue el costo económico, la privacidad y soberanía de datos. En muchos casos no fue posible usar herramientas web, ya que esto hubiera implicado el compartir los datos con una empresa y que estos se guardaran en un servidor extranjero. A esto se aunó el costo tanto del almacenamiento, ya fuera en discos o servidores, y del uso de herramientas propietarias. El resultado fue que se almacenaron los documentos en discos duros bajo el control de la CNB y se usaron los servidores del CIMAT para procesar conjuntos específicos de datos.

Criba

El proceso de criba consistió en seleccionar y priorizar documentos altamente relevantes (Angelus, 2022a). Los criterios de relevancia para la digitalización documental en el sistema Angelus se centraron en identificar documentos que aportaran indicios sobre la suerte o paradero de personas desaparecidas y sobrevivientes de desaparición durante las operaciones contrainsurgentes en México. Se priorizaron documentos producidos o resguardados por perpetradores (Sedena, DFS, DGIPS, etc.) o instituciones relacionadas (cuerpos de



seguridad, instituciones penitenciarias, de salud o asistencia) que contuvieran información sobre persecuciones, capturas, cautiverios, traslados, nacimientos en cautiverio, apropiaciones de menores, centros de detención, participación de perpetradores, operaciones represivas, y expedientes vinculados a la búsqueda, denuncia o reparación de víctimas. Asimismo, se incluyeron documentos integrados a expedientes administrativos o judiciales relacionados con la investigación de violaciones a derechos humanos, juicios de amparo, procesos penales o procedimientos de reparación. También se consideraron relevantes las fuentes publicadas, como prensa, literatura académica, testimonios, memorias o tesis, que investigan prácticas represivas, desapariciones forzadas, lugares de detención clandestina, *modus operandi* e identidad de perpetradores, o la estructura de agencias contrainsurgentes.

Esta selección fue necesaria debido a la enorme cantidad de documentos generados durante las operaciones de contrainsurgencia y esfuerzos posteriores de esclarecimiento, que asciende a decenas de millones de fojas. Esta gran masa documental contiene información heterogénea, fragmentaria y, en muchos casos, repetitiva, ya que la estructura administrativa de los expedientes frecuentemente provoca duplicación de documentos o inclusión de copias parciales en múltiples archivos. Dado que los recursos humanos y técnicos son limitados, los criterios de selección permitieron enfocar los esfuerzos en aquellos documentos que contenían información relevante para el esclarecimiento de desapariciones forzadas y otras violaciones a los derechos humanos. Además, una criba adecuada reduce el riesgo de sobrecargar los sistemas de análisis con información irrelevante, mejora la calidad de los datos generados y permite una utilización más eficiente de las herramientas de búsqueda, anotación y construcción del grafo de conocimiento.

Anotación

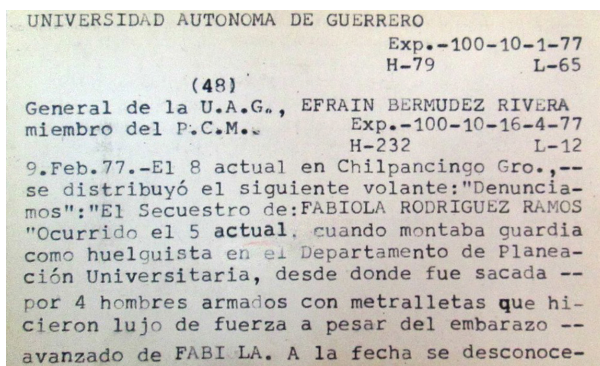
El proceso de anotación en el sistema Angelus constituye el núcleo metodológico que permite transformar la información documental histórica en conocimiento estructurado y consultable (Angelus, 2022b). Este procedimiento complejo involucra dos componentes fundamentales: la carga de documentos al repositorio digital y su posterior anotación manual utilizando la herramienta especializada denominada “Anotador”. El proceso requiere que usuarios expertos revisen documento por documento, identificando y registrando entidades relevantes como personas, lugares, eventos y organizaciones, así como las relaciones que existen entre ellas. Esta información se estructura según una ontología computacional preestablecida que permite su almacenamiento en forma de grafo de

conocimiento. La anotación requirió tanto la capacitación en el uso del Anotador, como el uso de criterios específicos que fueron desarrollados y refinados a lo largo del proyecto y una lectura contextual de los documentos.

El trabajo se estructuró de varias formas, para finalmente organizarse en equipos de anotadores, constituidos por múltiples becarios y coordinados por personal especializado de la CNB. Cada equipo fue encargado de revisar un conjunto específico de documentos previamente seleccionados por su relevancia. Esta organización por lotes permitió distribuir la carga de trabajo y dar seguimiento al avance del proyecto de manera sistemática. Se establecieron ciclos de revisión en los que las anotaciones realizadas eran validadas por personal con mayor experiencia, y se llevaron a cabo sesiones periódicas de retroalimentación colectiva para resolver dudas metodológicas o técnicas. A lo largo del proyecto se elaboraron guías actualizadas y ejemplos anotados (Angelus, 2022b), con el fin de reforzar los criterios y asegurar la consistencia entre los distintos participantes.

Un criterio fundamental que guía todo el proceso de anotación establece que cada enunciado debe reflejar fielmente la información contenida en el soporte documental anotado, sin realizar juicios sobre la veracidad de la fuente. Esta aproximación metodológica reconoce que los documentos históricos pueden contener errores, omisiones o perspectivas sesgadas, pero preserva la integridad del registro documental como fuente primaria. Por ejemplo, cuando una ficha de la DFS relata que estudiantes de la Universidad Autónoma de Guerrero denunciaron el secuestro de Fabiola Rodríguez Ramos el 5 de octubre de 1977 por hombres armados (figura 4), el proceso de anotación registra la captura de Fabiola por hombres armados y referencia la ficha original, permitiendo que los analistas posteriores puedan realizar sus propios juicios sobre la información proporcionada.

Figura 4. Ficha de la DFS sobre volantes distribuidos en Universidad Autónoma de Guerrero (9/feb/77)



Las dificultades del proceso de anotación son múltiples y complejas. El volumen y la heterogeneidad de los documentos impusieron límites operativos. Al 30 de noviembre de 2023, el sistema contabilizaba 10 743 entidades individuales, 31 599 propiedades sobre esas entidades y 48 781 relaciones entre entidades, cifras que reflejan tanto la magnitud del trabajo realizado como la necesidad de mantener procesos eficientes y consistentes a gran escala. La necesidad de mantener la trazabilidad documental de entidades y relaciones, y de mantener la coherencia entre anotaciones en el uso de nombres en clave, entidades duplicadas, variantes gráficas y eventos ambiguos requirió la reificación del grafo y la asignación de identificador de recursos uniformes (URI). Además, se integraron herramientas de verificación semiautomatizada y procesos de deduplicación (Ioannou *et al.*, 2010). Sin embargo, la existencia de múltiples versiones de un mismo hecho o formas de referirse a una misma persona, implicó que las decisiones siguieron dependiendo del criterio experto, por lo que fue necesario formar a los anotadores en la representación paralela de versiones contradictorias, cada una debidamente referida a su fuente documental.

El uso de lenguaje codificado o eufemístico presente en muchos documentos oficiales es otra dificultad del proceso, los anotadores deben interpretar términos como “paquetes” para referirse a personas capturadas e “interrogatorio intenso” para describir procesos de tortura. Otra dificultad significativa surge del anonimato de muchas víctimas mencionadas en la documentación. Frecuentemente, los documentos refieren a personas sin proporcionar nombres específicos, utilizando únicamente descripciones numéricas como “10 paquetes” o referencias vagas. Esta situación complica no solo la identificación de las víctimas, sino también su anotación en el sistema. Sin embargo, el proyecto ha desarrollado estrategias para abordar esta problemática, utilizando información contextual como fechas y lugares para diferenciar entidades anónimas similares. Por ejemplo, se recomienda nombrar a estas entidades como “10 personas detenidas en Atoyac de Álvarez el 10/05/1978”, proporcionando descriptores mínimos que permitan su identificación posterior.

La capacitación de los anotadores representó uno de los desafíos más complejos del proyecto, derivado tanto de la naturaleza técnica del sistema como de la complejidad metodológica requerida para el trabajo con documentos históricos sensibles. El proceso de capacitación debía cubrir múltiples dimensiones: el manejo técnico de la herramienta Anotador, la comprensión de la ontología computacional, la aplicación de criterios específicos de anotación, y el desarrollo de competencias archivísticas e históricas necesarias para interpretar documentación oficial del periodo de la guerra sucia. La curva de aprendizaje se vio significativamente afectada por los problemas administrativos que implicaron

periodos cortos de anotación separados por meses, la generación de ineficiencias por la pérdida de becarios, y la necesidad de capacitar y recapacitar a los becarios. Estas dificultades resaltan la importancia de la estabilidad laboral y emocional en la realización de cualquier trabajo.

El proceso de anotación debe responder a las necesidades específicas de búsqueda de personas desaparecidas, lo que significa que trasciende el ejercicio académico de organización de información para convertirse en un trabajo con implicaciones directas para las familias de las víctimas y las acciones gubernamentales. Las sesiones de capacitación y revisión no solo buscaron transmitir conocimientos técnicos, sino también concientizar a los participantes sobre la dimensión política, social y humana del trabajo. Cada anotación representa no solo un dato, sino también una pieza potencialmente clave en el esclarecimiento de hechos que afectaron a miles de personas. Esta responsabilidad impregnó de sentido ético al proceso de formación y consolidación del equipo de trabajo.

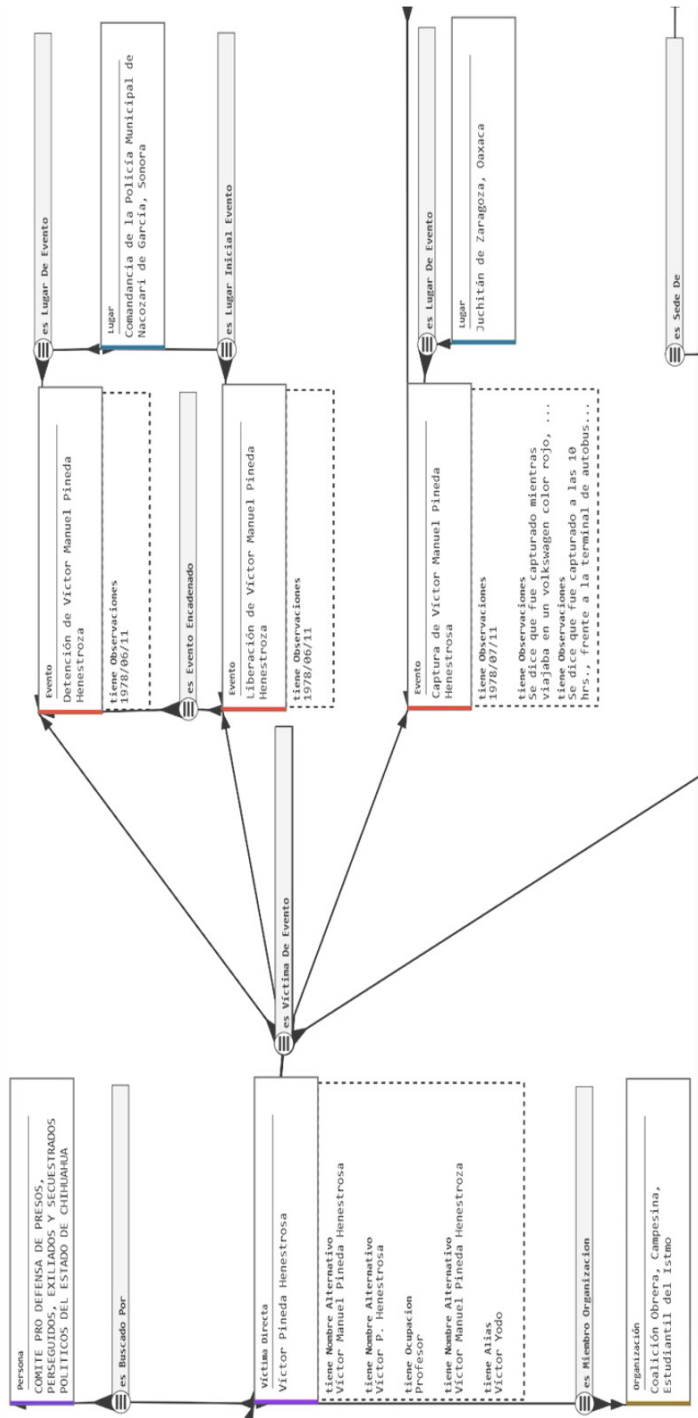
Visualización

El visualizador del sistema Angelus ha sido una herramienta clave para la revisión, análisis y representación de la información derivada de los documentos anotados (Casillas Pérez, 2023). Su desarrollo respondió a la necesidad de ofrecer una interfaz comprensible para diversos tipos de usuarios, desde investigadores hasta funcionarios públicos y familiares de personas desaparecidas. En una primera etapa, su principal uso fue facilitar la revisión de la anotación documental; sin embargo, conforme avanzó el proyecto y el grafo creció, fue posible realizar análisis más complejos, construir grafos temáticos y generar reportes automatizados.

Una de las virtudes del visualizador fue la adecuada selección de colores y direcciones para representar entidades y relaciones dentro del grafo. Esta elección contribuyó a que los usuarios pudieran identificar patrones o relaciones clave sin requerir conocimientos técnicos avanzados. Aunado a esto, la simbología y los colores facilitaron también la entrega de la información con las familias. Una práctica habitual durante la toma de testimonios a familiares y sobrevivientes era compartir con ellas la información con la que contaba el sistema hasta ese momento. Gracias a estas herramientas visuales, las personas que recibían los grafos comprendían fácilmente lo que significaba esa información. Incluso, en algunos casos, se hicieron correcciones sobre datos que mencionaban algunos documentos y que estaban registrados en Angelus.



Figura 5. Fragmento del grafo sobre la detención-desaparición de Víctor Pineda Henestroza



Un ejemplo de esto ocurrió con un grafo sobre la detención-desaparición de Víctor Pineda Henestrosa, quien fue detenido el 18 de julio de 1978 en Juchitán, Oaxaca (figura 5). Al momento de compartir el grafo con su hija Irma Pineda Santiago, ella aclaró que el nombre de su papá no era “Víctor Manuel Pineda Henestrosa”, como lo mencionaba un documento, así mismo, aclaró que un evento en donde se decía que Víctor fue detenido y liberado casi inmediatamente en Sonora era falso. No obstante, para Irma resultó relevante compilar en un mismo grafo las múltiples versiones que el Estado tenía sobre el paradero de su padre únicamente con el objetivo de desresponsabilizarse por la desaparición de Víctor.

Por otro lado, la forma confusa de representar el tiempo en la ontología y la imposibilidad de fijar fechas específicas generó dificultades, afectando especialmente aquellos casos en los que la secuencia de eventos es crítica para la reconstrucción de los circuitos de desaparición. La incorporación de capacidades adicionales, como la generación de reportes que integran todos los documentos asociados a una entidad o evento particular, implicó una mejora significativa en la navegación de los documentos.

A medida que se ampliaron los usos del sistema, el visualizador fue adoptado para tareas más allá de la verificación de datos. Se volvió una herramienta activa de investigación, al permitir explorar las conexiones entre personas, lugares y eventos, así como identificar elementos relevantes para la planificación de acciones de búsqueda. Sin embargo, este proceso también evidenció desafíos técnicos y de diseño, ya que el crecimiento del grafo y la variedad de casos complicaban la navegación. Para facilitar el entendimiento, los analistas realizaron grafos temáticos, los cuales fueron de gran utilidad para los familiares y las acciones de búsqueda, tal como se mencionó anteriormente en relación con el caso de Víctor Pineda.

Análisis

El sistema Angelus desempeñó un papel central en el proceso de análisis de los datos recopilados sobre desapariciones forzadas durante la guerra sucia, facilitando no solo la organización de grandes volúmenes de información documental, sino también su traducción en acciones concretas de búsqueda, generación de conocimiento histórico y vinculación con familiares de víctimas. Su estructura basada en grafos de conocimiento permitió integrar datos provenientes de fuentes heterogéneas (documentos oficiales, entrevistas, archivos personales) y transformarlos en entidades, relaciones y patrones analíticos explotables por distintos perfiles de usuario.



El sistema Angelus permitió a la Comisión Nacional de Búsqueda pasar de una lógica de investigación individualizada a una basada en la sistematización y el análisis relacional, de hecho establecida en el Protocolo Homologado de Búsqueda: búsqueda por patrones. A finales de 2023, el sistema registraba 2770 víctimas, 1240 presuntas personas perpetradoras, 6631 instituciones involucradas (incluyendo órganos de seguridad y organizaciones guerrilleras) y ocho entidades georreferenciadas del país donde ocurrieron los eventos (Guerrero, Ciudad de México, Morelos, Oaxaca, Veracruz, Michoacán, Estado de México y Puebla). Esta base acumulativa facilitó la elaboración de expedientes, informes y planes de acción, reduciendo significativamente los tiempos de procesamiento en comparación con los métodos tradicionales. Además, los grafos de conocimiento derivados del sistema fueron presentados ante fiscalías, organismos internacionales, colectivos y familiares, reforzando su utilidad en procesos de esclarecimiento y justicia.

Los productos generados con el sistema Angelus incluyeron la elaboración de una lista de lugares de interés vinculados con eventos de desaparición, la cual se generó mediante consultas estructuradas en el grafo; el diseño de entrevistas a familiares de seguimiento que retomaban testimonios previos, lo que ayudó a evitar su revictimización; boletines mensuales dirigidos a familiares documentando avances temáticos como: los “vuelos de la muerte”, las fosas comunes, o inspecciones en instalaciones militares; y el avance de la descripción del archivo histórico de la FEMOSPP, entre otros. La herramienta también impulsó una colaboración estratégica con instituciones como la Comisión Verdad y Justicia (CoVEH) y permitió la donación de insumos ya procesados a otras instancias investigativas.

El análisis de datos realizado con el sistema Angelus estuvo acompañado de una reflexión crítica sobre su origen y limitaciones. Es importante reconocer que la mayoría de los datos fueron generados en contextos de persecución estatal, lo que implicó asumir sus sesgos inherentes. Las decisiones de digitalización y anotación influyeron en la información contenida en el grafo y en la narrativa resultante. Una pregunta constante fue ¿cuáles son los vacíos del grafo?, ya que estas ausencias invisibilizan muchas veces a las poblaciones más vulnerables y evitan entender el problema y evitar su repetición. En conjunto, el sistema Angelus no solo funcionó como una herramienta técnica, sino también como un dispositivo político y epistemológico, al permitir construir conocimiento con incidencia directa en las estrategias de búsqueda, el acompañamiento a víctimas y la recuperación de la memoria histórica.

Conclusión

El desarrollo y uso del sistema Angelus ha sido, ante todo, un proceso de aprendizaje colectivo. Enfrentamos dificultades técnicas, metodológicas y humanas que no pueden ignorarse ni repetirse: la falta de herramientas existentes adaptadas al problema, la complejidad del manejo de datos contradictorios o inciertos, y los retos de capacitar a un equipo diverso en un entorno inestable. Compartir estas lecciones es esencial para que futuros proyectos no repitan los mismos errores y puedan avanzar más rápido, con mayor claridad y sensibilidad.

Los sistemas computacionales aplicados a los derechos humanos enfrentan límites estructurales que condicionan su utilidad y, en ocasiones, amplifican riesgos existentes. Su carácter de doble uso permite que las mismas herramientas empleadas para documentar violaciones puedan convertirse en instrumentos de vigilancia masiva, censura y represión estatal, comprometiendo la autonomía y seguridad de defensores y víctimas. Además, la presencia de sesgos en los datos y en los modelos reproduce desigualdades estructurales, especialmente en sistemas de alto riesgo como los usados en justicia, migración o verificación biométrica, donde los errores pueden traducirse en afectaciones graves a derechos fundamentales. A ello se suman déficits de gobernanza y transparencia que dificultan la rendición de cuentas, así como limitaciones metodológicas derivadas de la naturaleza opaca de muchos modelos computacionales y los datos sensibles, fragmentarios y difíciles de verificar en contextos de violencia. Estas tecnologías también generan dependencias técnicas, exclusiones lingüísticas y costos ambientales que deben ser considerados. En conjunto, estos límites subrayan que la tecnología es solo un apoyo: las decisiones críticas requieren juicio humano y marcos éticos sólidos que mantengan la centralidad de las personas y sus derechos.

Este esfuerzo nos obligó a cuestionar de forma constante: ¿ciencia de datos por y para quién? El sistema Angelus no fue pensado como una herramienta para satisfacer la curiosidad académica ni para producir métricas vacías. Desde su diseño, puso en el centro a las víctimas y a sus familias, priorizando la utilidad práctica y la responsabilidad ética. La ciencia de datos aplicada a violaciones graves de derechos humanos no puede ser neutral: debe responder a una vocación de justicia, memoria y reparación. Finalmente, es crucial subrayar que el sistema Angelus no es una máquina de la verdad. Su potencia reside en la sistematización y el análisis relacional, pero sus resultados dependen de la calidad y el origen de los datos, del criterio humano en la anotación, y de las decisiones tomadas en cada etapa del proceso. No sustituye al juicio experto ni a la experiencia humana. Es una herramienta para apoyar el esclarecimiento, no para imponer una versión única de los hechos. Su valor está en permitirnos ver más, dudar mejor y actuar con mayor fundamento.



Agradecimientos

A Víctor Mireles, por la revisión del texto, y a las personas que participaron en la coordinación, desarrollo, anotación y análisis del sistema Angelus: Aarón Andrade Valdivia, Aldo Osorio Hernández, Alejandra Blanco Villarreal, Amelia Sarai Abud López, Amílcar Nevárez, Analucina Garza González, Andrea Torrealba Torre, Balam Palma, Berenice Pedroza, Brenda González, Candela Bastos Camus, Carlos Inclán, Cindy López, Claudia Alejandra Pacheco Flores, Cuitláhuac Alfonso Galaviz Miranda, Daniel Efraín Navarro Granados, David Israel Aguilera Briones, Diego Antonio Franco de los Reyes, Diego Emiliano Bautista Páez, Donnet Esquivel Romero, Donnovan Romero, Eduardo Herrera García, Efraín Navarro, Elena Jaloma Pérez, Elizabeth Nava, Fernanda Espinosa Moreno, Fernanda Murillo, Fernando Cedeño, Francisco Casillas, Fátima Granillo Hernández, George Mitl, Gerardo Sánchez Nateras, Grecia Pérez, Hector Bravo, Héctor Amílcar Nevárez Fernández, Ileana García Rodríguez, Javier Perea Salguero, Javier Yankelevich Winocur, Jazmine Dafne Somellera Carrasco, Jessica Sánchez Rafael, Jorge Luis Vargas, Juan Camilo Pantoja García, Juan Carlos Gonzáles Aguilar, Juana Marisol Garzón Rivera, Karla Aguilar Díaz, Leticia Tecla Figueroa, Lucero San Vicente Juambelz, Luis Donnovan Romero Ángeles, Marcelo Alberto Sánchez, Marina Leticia Medel Pérez, Mario López Pérez, Martín González Romero, Martín H. González, María Guadalupe De Sampedro Paredes, Mónica Elizabeth Acevedo Gallegos, Norma Verónica Trinidad, Nuria Ivonne Hernández Hernández, Omar Villarreal Salas, Oscar De la Cruz Echeveste, Oscar Emmanuel Martínez, Otto Hahn Herrera, Ramón Alfredo Gracida Gómez, Rubén Ricardo Ponce Oviedo, Saúl Rivera Juárez, Sebastián Casillas, Túpac Amaru Gutiérrez Ortega, Verónica Vázquez Mata, Victor Muñiz, Viviana Beltrán, Víctor Mireles, Zacnité Chacón Estrada, Zacnité Minor.

Asimismo, al Conahcyt por el financiamiento a través del proyecto número 321368 que hizo posible desarrollar y operar la tecnología que sustenta estos esfuerzos, la cual ha sido puesta a disposición del público como *software* libre en <https://seguridad.conahcyt.mx/guerra-sucia/archivo>

Se usó AI generativa (Language Tool) para la revisión del texto.

Referencias

Aguayo, S. (2001). *La charola: una historia de los servicios de inteligencia en México*. México: Grijalbo.

- Angelus, E. (2022a). Proyecto angelus: Primer catálogo y reporte de digitalización. *Zenodo*. Recuperado de <https://doi.org/10.5281/zenodo.7548893>
- Angelus, E. (2022b). Proyecto Angelus: Primer reporte de anotación. *Zenodo*. Recuperado de <https://doi.org/10.5281/zenodo.7552704>
- Apache Jena Project. (2025). Fuseki 2 - sparql server. Recuperado el 29 de julio de 2025, de <https://jena.apache.org/documentation/fuseki2/>
- Archivo General de la Nación (AGN). (2020). *Composición del fondo DFS-DGIPS*. Recuperado el 29 de julio de 2025, de <https://www.gob.mx/agn/es/articulos/composicion-del-fondo-documental-dfs-dgips-estado-de-organizacion?idiom=es>
- Archivo Nacional de la Memoria (ANM). (s. f.). Presidenta de la Nación. Recuperado el 18 de noviembre de 2025, de <https://www.argentina.gob.ar/derechoshumanos/anm>
- Artículo 19. (2021). Archivos de la resistencia. Recuperado el 29 de julio de 2025, de <https://archivosdelaresistencia.org/>
- Artículo 19. (2015). Memoria y verdad. Archivos de la represión. Recuperado el 29 de julio de 2025, de <https://archivosdelarepresion.org/>
- Ballinas, V. (2001, 31 de julio). El entonces procurador general revela información sobre desaparecidos en guerrero durante la “guerra sucia”. *La Jornada*. Recuperado el 31 de julio de 2025, de <https://www.jornada.com.mx/2001/07/31/006n1pol.html>
- Blaser, M. (2013). Ontological conflicts and the stories of peoples in spite of Europe: Toward a conversation on political ontology. *Current Anthropology*, 54(5), 547-568. <https://doi.org/10.1086/672270>
- Calderón, F. H., & Cedillo, A. (2012). *Challenging authoritarianism in México*. Londres: Routledge. <https://doi.org/10.4324/9780203133224>
- Caridi, I., Dorso, C. O., Gallo, P., & Somigliana, C. (2011). A framework to approach problems of forensic anthropology using complex networks. *Physica A: Statistical Mechanics and its Applications*, 390(9), 1662-1676. <https://doi.org/10.1016/j.physa.2010.11.042>
- Carrington, P. J., Scott, J., & Wasserman, S. (2005). *Models and methods in social network analysis*. Vol. 28. Cambridge University Press. <https://doi.org/10.1017/CBO9780511811395>
- Casillas Perez, S. (2023). *Human-centered design and implementation of an ontology-based databank tool*. Tesis de maestría, Aalot University. <https://urn.fi/URN:NBN:fi:aalto-202305213350>



- Cedillo, A., & Valenzuela, F. C. (2014). Análisis de la producción historiográfica en torno a la “guerra sucia” mexicana. En V. Oikión Solano, E. Rey Tristán, & M. López Ávalos (Eds.), *El estudio de las luchas revolucionarias en América Latina (1959-1996): Estado de la cuestión* (pp. 263-288). México: El Colegio de Michoacán/Universidad de Santiago de Compostela.
- Comisión de la Verdad del Estado de Guerrero (ComVerdad). (2014). *Informe final de actividades*. Recuperado el 29 de julio de 2025, de <http://congresogro.gob.mx/files/InformeFinalCOMVERDAD.pdf>
- Comisión Nacional de Derechos Humanos. (2001). Recomendación 26/2001. CNDH México. Recuperado el 29 de julio de 2025, de <https://www.cndh.org.mx/documento/recomendacion-262001>
- Consejo Nacional de Humanidades, Ciencias y Tecnologías (Conahcyt). (2024a). *Angelus — guerra sucia*. Recuperado el 29 de julio de 2025, de <https://seguridad.conahcyt.mx/guerra-sucia/angelus>
- Consejo Nacional de Humanidades, Ciencias y Tecnologías (Conahcyt). (2024b, 19 de junio). *Conahcyt y CNB celebran transferencia tecnológica de Angelus, software para la búsqueda de personas desaparecidas y no localizadas*. Recuperado el 29 de julio de 2025, de <https://www.gob.mx/segob/prensa/apoyara-conahcyt-a-cnb-para-busqueda-de-personas-desaparecidas-y-no-localizadas>
- D’ignazio, C., & Klein, L. F. (2023). *Data feminism*. MIT Press.
- De la Cruz Echeveste, O. (2023). *Extracción y análisis de texto en expedientes relacionados con desapariciones forzadas durante el periodo de la “guerra sucia” en México*. Tesis de Maestría en Cómputo Estadístico, Centro de Investigación en Matemáticas, Unidad Monterrey, Monterrey, Nuevo León.
- Dirección Federal de Seguridad (DFS). (1971). *Expediente de la dirección federal de seguridad (DFS) con el seguimiento de 314 militantes a nivel nacional*. Biblioteca digital Archivos de la Represión. Artículo 19/AGN. Recurso digital, cubre el periodo de junio 1971 a agosto 1978. Recuperado el 31 de julio de 2025, de <https://biblioteca.archivosdelarepresion.org/item/91005>
- Dueck, J., Guzman, M., & Verstappen, B. (2001). *HURIDOCS Events Standard Formats: A tool for documenting human rights violations*. (2nd rev. ed.). Versoix, Suiza: HURIDOCS. Recuperado de https://huridocs.org/wp-content/uploads/2020/11/HURIDOCS_ESF_English1.pdf
- Fiscalía Especial para Movimientos Sociales y Políticos del Pasado (FEMOSPP). (2006). *Informe histórico presentado a la sociedad mexicana*. National Security Archive. Recuperado el 29 de julio de 2025, de <https://nsarchive2.gwu.edu/NSAEBB/NSAEBB209/informe/intro.pdf>

- Gamiño Muñoz, R. (2020). Los orígenes de la “verdad histórica”. Los primeros informes sobre la desaparición forzada de personas en México. *Relaciones. Estudios de Historia y Sociedad*, 41(160), 1-25. <https://doi.org/10.24901/rehs.v41i161.662>
- González Aguilar, J. C. (2022). *Uso de aprendizaje de máquina para el análisis de los ficheros de la dirección federal de seguridad (DFS)*. Tesis de Licenciatura en Matemáticas Aplicadas y Computación, Facultad de Estudios Superiores Acatlán, Universidad Nacional Autónoma de México.
- Green, A. H., & Ball, P. (2019). Civilian killings and disappearances during civil war in El Salvador (1980-1992). *Demographic Research*, 41, 781-814. <https://doi.org/10.4054/DemRes.2019.41.27>
- Guberek, T., & Hedstrom, M. (2017). On or off the record? Detecting patterns of silence about death in Guatemala's National Police Archive. *Archival Science*, 17, 27-54. <https://doi.org/10.1007/s10502-017-9274-3>
- HURIDOCs Team. (2025). *Using machine learning in Uwazi to support human rights documentation work*. HURIDOCs. Recuperado de <https://huridocs.org/2025/04/using-machine-learning-in-uwazi-to-support-human-rights-documentation-work/>
- Ioannou, E., Papapetrou, O., Skoutas, D., & Nejdil, W. (2010). Efficient semantic-aware detection of near duplicate resources. En L. Aroyo *et al.* (Eds.), *The semantic web: Research and applications* (Vol. 6088, pp. 136-150). Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-642-13486-9>
- José López Portillo. (1978). Ley de Amnistía (abrogada, D.O.F. 28 de septiembre de 1978). PDF en sitio oficial de la Cámara de Diputados. Recuperado el 31 de julio de 2025, de <https://www.diputados.gob.mx/LeyesBiblio/abro/lamn78/LAmn78.abro.pdf>
- Koho, M., Ikkala, E., Leskinen, P., Tamper, M., Tuominen, J., & Hyvönen, E. (2020). WarSampo knowledge graph: Finland in the Second World War as Linked Open Data. *Semantic Web: – Interoperability, Usability, Applicability*, 12(2), 265-278. <https://doi.org/10.3233/SW-200392>
- Lanthaler, M., Wood, D., & Cyganiak, R. (2014). Rdf 1.1 concepts and abstract syntax. W3C, W3C Recommendation. <https://www.w3.org/TR/rdf11-concept>
- Levin, L., & Hamber, B. (1997). *The Use of evsys for Preparing a Human Rights Database for Presentation to the Truth and Reconciliation Commission (TRC) in South Africa*. HURIDOCs Conference, Centre for the Study of Violence and Reconciliation. <https://www.csvr.org.za/the-use-of-evsys-for-preparing-a-human-rights-database-for-presentation-to-the-truth-and-reconciliation-commission-trc-in-south-africa/>



- Mecanismo para el Acceso a la Verdad y el Esclarecimiento Histórico (MEH). (2024a). *Fue el Estado (1965-1990). Presentación, resumen y recomendaciones del informe final del mecanismo para la verdad y el esclarecimiento histórico*. (Vol. 1: A. B. Hernández, D. F. Dávalos, & C. A. P. Ricart, Eds.). Ciudad de México: Secretaría de Gobernación. <https://www.meh.org.mx>
- Mecanismo para el Acceso a la Verdad y el Esclarecimiento Histórico (MEH). (2024b). *Fue el Estado (1965-1990). ¿Quiénes fueron los perpetradores? Instituciones y personas implicadas en la comisión de violaciones graves a derechos humanos (1965-1990)*. (Vol. 4; A. B. Hernández, D. F. Dávalos, & C. A. P. Ricart, Eds.). Ciudad de México: Secretaría de Gobernación. <https://www.meh.org.mx>
- Mireles Chávez, V., Martínez Sánchez, M., Yankelevich Winocur, J., Sánchez Nateras, G., et al. (2021). Buscando a los desaparecidos de la guerra sucia: ontologías computacionales y la búsqueda de verdad. *Iberoforum*, 1(1), 1-40. <https://doi.org/10.48102/ibf.2021.v1.n1.149>
- Price, M., Guberek, T., Guzmán, D., Zador, P., & Shapiro, G. (2009). *A statistical analysis of the Guatemalan National Police Archive: Searching for documentation of human rights abuses*. Paper presented at the Joint Statistical Meetings, Section on Survey Research Methods. <https://hrdag.org/wp-content/uploads/2013/02/JSM-GT-estimates.pdf>
- Puerto, M. S., Abboud, D., Baraybar, J. P., Carracedo, A., Fonseca, S., Goodwin, W., et al. (2021). The search process: Integrating the investigation and identification of missing and unidentified persons. *Forensic Science International: Synergy*, 3, 100154. <https://doi.org/10.1016/j.fsisyn.2021.100154>
- Rodríguez Kuri, A. (2024). Mujeres en la clandestinidad armada. Hechos y tendencias en la década de 1970. *Historia Mexicana*, 74(1), 325-379. <https://doi.org/10.24201/hm.v74i1.4795>
- Santiago, V. (2022). *Angelus: el algoritmo que escarba en la guerra sucia*. Recuperado el 29 de julio de 2025, de <https://www.gatopardo.com/articulos/angelus-el-algoritmo-que-escarba-en-la-guerra-sucia?5b6d75f9> page=73 (Publicado originalmente en Gatopardo).
- Schneider, M., Carroll, J., Herman, I., Patel-Schneider, P. F., et al. (2009). Owl 2 web ontology language rdf-based semantics. W3C recommendation (October 27 2009).
- The Dataverse Project. (2025). *Dataverse*. Recuperado el 29 de julio de 2025, de <https://dataverse.org/>
- Ushahidi. (s. f.). *Ushahidi*. Recuperado el 18 de noviembre de 2025, de <https://www.ushahidi.com/>
- Vicente Ovalle, C. (2019). *Tiempo suspendido: una historia de la desaparición forzada en México, 1940-1980*. México: Bonilla Artigas.