

Análisis de Dominio e inteligencia artificial: una relectura contemporánea de Hjørland

Gustavo Liberatore

<https://orcid.org/0000-0002-5841-8739>

Resumen: el presente documento constituye un breve ensayo sobre la relación de la teoría del Análisis de Dominio de Hjørland y la Inteligencia Artificial a partir de los aportes más recientes desde el campo de la Organización del Conocimiento. Su objetivo es contribuir a la discusión y análisis, dentro de un entorno de aprendizaje, del impacto de estas tecnologías en la concepción más tradicional de este enfoque. Este texto se propone como material docente dentro de la asignatura de Organización del Conocimiento II de la UNMDP.

Cita sugerida: Liberatore, G. (2026). *Análisis de Dominio e inteligencia artificial: una relectura contemporánea de Hjørland* [Material docente no publicado, Universidad Nacional de Mar del Plata].

La teoría del Análisis de Dominio (AD) formulada por Birger Hjørland y Hanne Albrechtsen (1995) conserva una notable vigencia para pensar los problemas contemporáneos de la organización del conocimiento (OC), incluso en el actual contexto de expansión de la inteligencia artificial (IA). Lejos de haber quedado desplazado por el desarrollo de sistemas automáticos de procesamiento del lenguaje, el AD ofrece un marco especialmente fértil para comprender sus alcances, sus limitaciones y sus condiciones de validez. En su formulación, Hjørland y Albrechtsen sostienen que el objeto de análisis para la Ciencia de la Información no es el individuo aislado, sino los dominios del conocimiento entendidos como comunidades de pensamiento o de discurso, inscriptas en la división social del trabajo. Desde esta perspectiva, la OC, los sistemas de organización del conocimiento (SOCs), los criterios de pertinencia y las formas de comunicación no pueden comprenderse por fuera de los marcos sociales, históricos y epistémicos en los que se producen.

Esta tesis resulta particularmente significativa frente a la IA contemporánea. Los modelos actuales —y en especial los grandes modelos de lenguaje— operan a partir de regularidades estadísticas extraídas de corpus masivos, pero esa capacidad de procesamiento no equivale por sí misma a una comprensión epistemológica de los dominios. Precisamente por ello, el AD permite introducir una variable decisiva: el significado, la relevancia y la validez de la información no son propiedades abstractas ni universales, sino construcciones ligadas a comunidades discursivas específicas, a tradiciones de investigación, a controversias teóricas, a instituciones y a prácticas de validación (Romero Quesada, 2012). En este sentido, la IA no invalida el planteamiento teórico de Hjørland sobre el AD, sino que refuerza su

actualidad, porque muestra hasta qué punto la automatización semántica requiere un anclaje fuerte en estructuras de conocimiento socialmente constituidas.

No obstante, una actualización posible del AD consiste, entonces, en desplazar la pregunta desde “qué puede hacer la IA por sí sola” hacia “bajo qué condiciones de dominio la IA produce resultados conceptualmente válidos”. Desde esta perspectiva, la IA no debería ser pensada como sustituto del análisis humano del conocimiento, sino como una herramienta heurística e infraestructural cuyo rendimiento depende de la calidad epistemológica del dominio reconstruido. Esto supone reconocer que todo sistema basado en IA trabaja, explícita o implícitamente, sobre decisiones previas: qué corpus representan legítimamente un campo, qué conceptos se consideran centrales, qué relaciones semánticas se aceptan, qué voces se jerarquizan, qué sesgos se corrigen y qué criterios de autoridad se adoptan. La potencia técnica del modelo no elimina esas decisiones; por el contrario, las vuelve más opacas si no son sometidas a examen crítico. Esta orientación coincide con desarrollos recientes de Hjørland, quien insiste en que los sistemas de información deben orientarse a proveer información fiable y autorizada, priorizando la calidad, los valores y los fines declarados en los documentos antes que la mera adaptación a preferencias inmediatas de los usuarios.

Desde esta relectura, puede sostenerse que el dominio ya no debe concebirse sólo como comunidad discursiva en sentido clásico, sino también como una especie de *ecología sociotécnica*¹ dentro del entorno informativo. En la actualidad, los dominios producen no sólo artículos, libros, SOCs, sino también datasets, grafos de conocimiento, vocabularios interoperables, modelos de embeddings², sistemas de recuperación aumentada y flujos automatizados de descripción e indexación. Sin embargo, este agregado no obliga a abandonar los principios de Hjørland. Más bien exige ampliarlos: los artefactos técnicos deben ser interpretados como nuevas mediaciones de la producción y comunicación del conocimiento, sin perder de vista que su sentido sigue dependiendo de prácticas sociales, instituciones, tradiciones disciplinares y formas de legitimación. La técnica modifica la materialidad del dominio, pero no reemplaza su historicidad ni su estructura social.

En este punto, el texto de 2002 resulta especialmente anticipatorio, porque Hjørland incluye explícitamente entre los once enfoques del AD el estudio de la cognición profesional y la inteligencia artificial. Allí señala que existe un solapamiento entre el AD y la informática, sobre todo en cuestiones vinculadas con terminología, significado y diseño de sistemas que reflejen un dominio antes que meras estructuras mentales individuales. Aun así, subraya que la perspectiva de la Ciencia de la Información debe ser más amplia y no limitarse al desarrollo computacional, ya que necesita comprender disciplinas y grupos de usuarios desde marcos humanistas y sociológicos. Este señalamiento es decisivo para el presente: la IA puede integrarse al AD, pero sólo si se evita reducir el dominio a una familia de problemas técnicos y se preserva su carácter social y epistemológico.

¹ La *ecología sociotécnica* estudia la interconexión profunda entre los sistemas sociales (comunidades, instituciones) y los sistemas técnicos o ambientales, entendiendo que la crisis ecológica no es solo ambiental, sino que proviene de la estructura social.

² Los *modelos de embeddings* son herramientas de IA que convierten datos (texto, imágenes, audio) en vectores numéricos densos, capturando su significado semántico para medir similitudes.

La bibliografía reciente que aborda esta articulación (AD/IA), aunque todavía no es abundante, confirma esta línea de lectura. Un trabajo particularmente cercano a este problema es el de Thellefsen y Friedman (2025), que discute de manera directa si la inteligencia artificial puede enriquecer el AD. Su conclusión general no es tecnocéntrica: la integración de IA sólo resulta fructífera cuando los sistemas se diseñan con sensibilidad hacia la especificidad de los dominios, sus contextos culturales y disciplinares y la naturaleza social de la producción de significado. En esa misma dirección, el artículo remarca que la OC no debe fundarse en modelos universales de usuario, sino en realidades epistémicas situadas.

Otros trabajos recientes permiten reforzar esta discusión desde planos más operativos. Kausch (2025) propone pensar la traducción del conocimiento como método interdisciplinario en un contexto de AD que incorpora tecnologías de aprendizaje automático para mejorar la interoperabilidad entre SOCs. Allí se sugiere que *embeddings* y modelos lingüísticos pueden servir como herramientas para explorar contextos semánticos, siempre que su uso permanezca guiado por una reflexión teórica sobre los dominios. Roncaglia (2026) en una investigación sobre catalogación, metadatos e IA generativa argumenta que los LLMs pueden apoyar la creación de metadatos, pero sólo bajo estrategias de integración cautelosas, como el uso de conjuntos de datos controlados, sistemas especializados o recuperación aumentada con bases autorizadas; además, advierte sobre alucinaciones, sesgos y desajustes entre modelos probabilísticos y arquitecturas formales del conocimiento.

También en el terreno aplicado se observa un movimiento convergente. Existen investigaciones recientes sobre enriquecimiento de ontologías mediante grandes modelos de lenguaje, indización temática automática basada en vocabularios controlados y desafíos compartidos de evaluación para clasificación de materias y aprendizaje ontológico (Kollapally et al., 2025). En conjunto, estos trabajos muestran que la IA ya está ingresando de lleno en tareas clásicas de la OC, como la extracción de términos, la sugerencia de conceptos, el alineamiento terminológico y la clasificación temática. Sin embargo, la misma literatura muestra que tales procesos siguen dependiendo de *ontologías semilla*³, vocabularios controlados, validación experta y criterios explícitos de evaluación, lo cual confirma una intuición central del AD: la OC no puede reducirse a correlaciones formales, porque requiere una interpretación sustantiva de los conceptos y de sus relaciones dentro de un campo.

Por todo ello, una resignificación contemporánea del AD podría formularse del siguiente modo: la IA no reemplaza el conocimiento de dominio, sino que lo presupone y lo vuelve más visible como problema teórico y metodológico. Cuanto más sofisticados son los sistemas de IA, más necesario resulta preguntarse qué dominio representan, desde qué corpus lo modelizan, qué jerarquías semánticas cristalizan, qué perspectivas excluyen y bajo qué criterios consideran “relevante”, “correcta” o “autoritaria” una respuesta. En este marco, el AD aporta justamente lo que los enfoques puramente técnicos no pueden garantizar por sí

³ Una ontología semilla es un núcleo mínimo o estructura básica de conceptos y relaciones, diseñado para iniciar el desarrollo de ontologías más completas y complejas. Sirve como punto de partida para modelar un dominio específico, permitiendo su expansión y formalización, a menudo apoyándose en inteligencia artificial.

solos: una teoría social del significado, una epistemología de la pertinencia y una crítica de los supuestos que subyacen a toda arquitectura de OC.

Leer hoy a Hjørland desde la perspectiva de la IA no implica modernizar superficialmente una teoría previa, sino reconocer que sus tesis fundamentales se vuelven aún más pertinentes en un entorno atravesado por automatizaciones semánticas. La lección principal del AD sigue siendo válida: los SOC's deben construirse a partir de la comprensión de comunidades discursivas concretas, de sus lenguajes, de sus tradiciones y de sus conflictos. La IA puede ampliar la escala, la velocidad y ciertas capacidades de exploración del trabajo documental; pero sin los enfoques propuestos por el AD, corre el riesgo de organizar el conocimiento de manera eficaz en términos computacionales y pobre en términos epistemológicos.

Hay una manera posible de actualizar el AD desde la IA sin alejarse o desechar a Hjørland: no pensar la IA como sustituto del dominio, sino como un artefacto que sólo funciona bien cuando el dominio ya fue epistemológica, social e institucionalmente reconstruido. En sus dos textos base, Hjørland (1995, 2002) sitúa el objeto de estudio del AD en las comunidades de pensamiento o discurso, insertas en la división social del trabajo. Además, afirma que los sistemas de información, el lenguaje, la cooperación y los criterios de pertinencia reflejan los objetos de trabajo de esas comunidades, no sólo estados mentales individuales.

Una posible reflexión con base en estos argumentos es que la IA no vuelve obsoleto al AD; lo vuelve más necesario. Los modelos actuales operan sobre grandes regularidades estadísticas del lenguaje, pero justamente por eso tienden a generalizar, homogeneizar y “promediar” tradiciones epistémicas diferentes. El AD ofrece, en este sentido, un ajuste: recordar que el significado, la relevancia y la validez informacional no son universales ni neutros, sino dependientes de teorías, instituciones, controversias y usos situados dentro de un dominio. Hjørland mismo insistirá más tarde (2024) en que los criterios de relevancia están implicados por las teorías vigentes en un dominio y que el conocimiento debe hacerse explícito en la metodología de la OC.

Dicho de otro modo: desde una lectura contemporánea, un LLM no “comprende” un dominio en sentido fuerte; reconstruye patrones discursivos de ese dominio a partir de corpus, etiquetas, ontologías, metadatos y usos previos. Eso no invalida su utilidad, pero obliga a desplazar su sentido: la IA debe ser entendida como instrumento heurístico e infraestructural, no como autoridad epistemológica final. Esta forma de leer la IA es muy afín al espíritu de Hjørland, porque mantiene la primacía del contexto social, de la historicidad de los conceptos y de la especificidad disciplinar. A la vez, conecta con trabajos recientes que advierten que la IA puede beneficiar al AD sólo si se integra cuidando su profundidad epistemológica y evitando interpretaciones que diluyan los matices propios de cada campo.

Desde una revisión de la bibliografía específica a partir de los textos canónicos de Hjørland sobre el AD, no parece todavía consolidarse como un subcampo autónomo el estudio de la relación e implicancias del binomio AD + IA, aunque sí como un núcleo emergente de discusión.

A partir de estas consideraciones, pueden sugerirse cinco claves de resignificación del AD desde los actuales avances de la IA:

- Primera clave: *del dominio como comunidad discursiva al dominio como ecología sociotécnica.*
Hjørland no pierde vigencia si añadimos que hoy los dominios no sólo producen artículos y clasificaciones: también producen datasets, embeddings, prompts, pipelines de RAG⁴, grafos de conocimiento y sistemas automáticos de indexación. Todo eso forma parte del modo en que un campo estabiliza conceptos y decide qué cuenta como evidencia, autoridad o pertinencia. El dominio sigue siendo social; simplemente ahora está mediado por nuevas infraestructuras técnicas.
- Segunda clave: *la IA puede ampliar el análisis, pero no definir sola el dominio.*
Hjørland ya advertía que no todos los dominios son iguales y que no pueden tratarse con una lógica uniforme. Un dominio no está simplemente “dado”, sino que se constituye a través de perspectivas, valores, intereses y teorías. Por eso, cualquier clasificación automática, cluster temático o taxonomía sugerida por IA debe leerse como una hipótesis analítica sobre el dominio, no como su verdad definitiva.
- Tercera clave: *la IA es especialmente útil en tareas de descubrimiento provisional.*
Donde mejor encaja con el AD es en la generación de candidatos: extracción de términos, detección de relaciones conceptuales, enriquecimiento de ontologías, sugerencia de etiquetas, identificación de géneros documentales, mapeo de controversias, detección de proximidades bibliométricas o interdisciplinares. Hoy ya hay trabajos que muestran LLMs aplicados a aprendizaje y enriquecimiento de ontologías, así como a asignación automática de materias en catálogos. Pero en todos esos casos la estructura resultante necesita validación semántica y control de dominio.
- Cuarta clave: *en AD, la validación no es sólo técnica; es epistémica.*
Un dominio construido en base a IA puede contener una buena similitud estructural y, aún así, carecer de consistencia si excluye algunas corrientes de pensamiento, aplana conceptos polémicos o impone una jerarquía ajena a la tradición del campo. Éste es el punto más crítico de todos: la representación debe siempre transparentar qué teorías están presupuestas, qué actores quedan representados, qué instituciones legitiman esa representación y qué intereses sirve. Los trabajos recientes sobre IA y

⁴ Un pipeline de RAG (Generación Aumentada por Recuperación) es un flujo de trabajo de IA que mejora los modelos de lenguaje (LLMs) buscando información externa relevante antes de generar una respuesta. Consta de dos fases principales: ingesta (cargar, segmentar y vectorizar documentos en una base de datos) y recuperación/generación (buscar datos pertinentes para responder consultas), lo que garantiza respuestas precisas y actualizadas.

OC hacen hincapié justamente sobre riesgos de sesgo, opacidad, exclusión epistémica y pérdida de diversidad semántica.

- Quinta clave: *el AD ofrece el marco para mejorar la IA mediante conocimiento experto.*

Hjørland sostiene que los especialistas temáticos siguen siendo esenciales para clasificaciones, bases de datos de alto nivel y ontologías avanzadas. Eso hoy adquiere una fuerza renovada: la IA puede acelerar el trabajo clasificatorio, pero necesita corpus curados, vocabularios controlados, criterios de autoridad, revisión experta y, en muchos casos, estrategias como RAG apoyado en fuentes autorizadas. Incluso en catalogación y metadatos, la literatura reciente propone precisamente esa dirección: modelos especializados, recuperación aumentada con bases confiables y alineación con arquitecturas formales de conocimiento.

En síntesis, la IA no reemplaza el AD, sino que lo aumenta y mejora. Cuanto más potente es la automatización semántica, más indispensable se vuelve una teoría social y epistemológica del dominio que permita decidir qué corpus usar, qué conceptos privilegiar, qué relaciones aceptar, qué sesgos corregir y qué criterios de relevancia adoptar. En un enfoque dialógico entre la IA y el AD, la pregunta no es si la máquina puede organizar conocimiento, sino bajo qué teoría del dominio, con qué corpus, para qué comunidad, con qué criterios de pertinencia y bajo qué controles epistémicos y éticos puede realizarse.

Referencias

- Hjørland, B. (2002). Domain analysis in information science: Eleven approaches - traditional as well as innovative. *Journal of Documentation*, 58(4), 422-462. <https://doi.org/10.1108/00220410210431136>
- Hjørland, B. (2024). *Some core concepts of domain analysis*. *Em Questão*, 30. <https://doi.org/10.1590/1808-5245.30.141781>
- Hjørland, B., & Albrechtsen, H. (1995). *Toward a new horizon in information science: Domain-analysis*. *Journal of the American Society for Information Science*, 46(6), 400-425. [https://doi.org/10.1002/\(SICI\)1097-4571\(199507\)46:6<400::AID-ASIS2>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASIS2>3.0.CO;2-Y)
- Kausch, J. (2025). *Knowledge translation as an interdisciplinary method for information science*. *Journal of Documentation*, 81(3), 569-580. <https://doi.org/10.1108/JD-07-2024-0167>
- Kollapally, N. M., Geller, J., Keloth, V. K., He, Z., & Xu, J. (2025). *Ontology enrichment using a large language model: Applying lexical, semantic, and knowledge network-based similarity for concept placement*. *Journal of Biomedical Informatics*, 168, Article 104865. <https://doi.org/10.1016/j.jbi.2025.104865>
- Romero Quesada, M. A. (2012). El análisis de dominio y sus presupuestos teóricos-metodológicos. *Bibliotecas. Anales de Investigación*, 8-9(8-9), 228-238.
- Roncaglia, G. (2026). *Cataloguing, metadata, and generative AI. Early experiences and future perspectives*. *JLIS.it*, 17(1), 106-127. <https://doi.org/10.36253/jlis.it-693>

- Thellefsen, M., & Friedman, A. (2025). *From realism and socio-cognitivism to AI constructs: Enhancing domain analysis through artificial intelligence?* *Information Research*, 30(CoLIS), 628–642. <https://doi.org/10.47989/ir30CoLIS52339>