



ARTÍCULO

Corrientes filosóficas y toma de decisiones en sistemas inteligentes de automatización industrial: un marco de capacidades para la Industria 4.0

Philosophical currents and decision-making in intelligent industrial automation systems: a capability framework for Industry 4.0

Danny Caicedo ^{*} · Valmore Márquez · Blanca Liliana González Pertúz

Universidad Privada Dr. Rafael Belloso Chácín, Maracaibo, Venezuela

*Autor de Correspondencia: dcaicedo@urbe.edu.ve

Enviado: 13 de Mayo Aceptado: 12 de Julio Publicado en línea: 15 de Julio

Resumen

La incorporación de sistemas inteligentes en la automatización industrial transforma la supervisión y la toma de decisiones, exigiendo profesionales capaces de combinar dominio técnico, comprensión causal, interpretación contextual y responsabilidad ante sus consecuencias. En este contexto, el objetivo fue analizar la contribución del positivismo, el realismo crítico y la hermenéutica a la formación de capacidades y la toma de decisiones, para proponer un marco integrado orientado a la Industria 4.0. Se adoptó un enfoque transdisciplinario y una metodología filosófico-analítica documental, sustentada en una revisión de alcance realizada en Scopus, Web of Science y Google Scholar durante 2020–2026. Tras aplicar los criterios de selección, se conformó un corpus de 22 fuentes, analizado mediante codificación temática. Los resultados muestran que el positivismo sustenta la medición y el control, el realismo crítico amplía el diagnóstico mediante mecanismos causales y la hermenéutica incorpora la interpretación del contexto humano–tecnología. Esta articulación permitió construir un marco de capacidades organizado en niveles operativo, táctico y estratégico, con la filosofía ambiental como dimensión transversal y la supervisión humana ante decisiones mediadas por inteligencia artificial. Se concluye que esta integración podría fortalecer el juicio profesional frente a sistemas industriales progresivamente automatizados, aunque requiere validación posterior.

Palabras clave: Automatización industrial, sistemas inteligentes, toma de decisiones, positivismo, realismo crítico, hermenéutica

Abstract

The incorporation of intelligent systems into industrial automation is transforming supervision and decision-making, requiring professionals capable of combining technical expertise, causal understanding, contextual interpretation, and responsibility for the consequences of their decisions. In this context, the objective was to analyze the contribution of positivism, critical realism, and hermeneutics to capability development and decision-making in order to propose an integrated framework for Industry 4.0. A transdisciplinary approach and a documentary philosophical-analytical methodology were adopted, supported by a scoping review conducted in Scopus, Web of Science, and Google Scholar covering the period 2020–2026. After applying the selection criteria, a corpus of 22 sources was established and analyzed through thematic coding. The results show that positivism supports measurement and control, critical realism extends diagnosis through the identification of causal mechanisms, and hermeneutics incorporates the interpretation of the human–technology context. This articulation enabled the development of a capability framework organized into operational, tactical, and strategic levels, with environmental philosophy as a cross-cutting dimension and human oversight of decisions mediated by artificial intelligence. It is concluded that this integration could strengthen professional judgment in increasingly automated industrial systems, although further validation is required.

Keywords: Industrial automation, intelligent systems, decision-making, positivism, critical realism, hermeneutics

Este es un artículo de acceso abierto y se distribuye bajo los términos de la Licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional (CC BY-NC-ND 4.0).

Citar como: D. Caicedo, “Corrientes filosóficas y toma de decisiones en sistemas inteligentes de automatización industrial: un marco de capacidades para la Industria 4.0,” *CISAI*, vol. 2, no. 2, pp. 44–62, Julio 2026, doi: 10.64439/cisai.v2i2.44.

1. Introducción

La denominada Cuarta Revolución Industrial o Industria 4.0 ha impulsado una transformación sin precedentes en los entornos productivos, caracterizada por la hiperconectividad, los sistemas ciberfísicos y la analítica de datos a gran escala, dentro de un escenario en el que la automatización industrial ha dejado de limitarse a la sustitución de la intervención humana en tareas repetitivas mediante esquemas rígidos de control lógico programable para evolucionar hacia ecosistemas socio-técnicos en los que convergen dispositivos, datos, algoritmos y personas. Esta evolución puede observarse en las arquitecturas contemporáneas de automatización, que incorporan capacidades distribuidas de procesamiento, comunicación y decisión en diferentes niveles de los sistemas industriales [1], ampliando las exigencias asociadas con su operación y supervisión y desplazando progresivamente la atención hacia las capacidades requeridas para comprender las relaciones que se establecen entre los componentes tecnológicos y humanos.

En correspondencia con esta ampliación de las capacidades requeridas, la formación profesional se enfrenta a cuestiones que trascienden el dominio estrictamente técnico y alcanzan la manera en que los ingenieros observan, explican e interpretan el comportamiento de los sistemas, especialmente si se considera que la formación en ingeniería ha permanecido tradicionalmente vinculada a una visión técnico-instrumental, mientras los sistemas tecnológicos contemporáneos incorporan dimensiones ontológicas, epistemológicas y éticas que intervienen tanto en su diseño como en su utilización [2]. Esta relación adquiere mayor complejidad cuando los procesos de decisión incorporan sistemas inteligentes cuyas operaciones internas pueden dificultar la interpretación y la supervisión humana efectiva [3, 4], configurando escenarios en los que el conocimiento instrumental comienza a relacionarse con formas de razonamiento orientadas a explicar causalidades, interpretar contextos y valorar las implicaciones derivadas de las decisiones tecnológicas.

A partir de esta coexistencia entre conocimiento instrumental y formas diversas de razonamiento, distintas perspectivas filosóficas permiten aproximarse a dimensiones complementarias de la realidad tecnológica, entre ellas el positivismo lógico, vinculado con la observación empírica y la cuantificación, aunque su ideal de neutralidad valorativa ha sido cuestionado en la formación reciente en ingeniería [5], y el realismo crítico, que incorpora la retroducción causal para explorar mecanismos subyacentes que no siempre resultan directamente observables [6], mediante procedimientos que permiten estructurar este tipo de razonamiento [7]. A estas aproximaciones se suma la hermenéutica, desde la cual adquieren relevancia la interpretación contextual y las relaciones entre tecnología y experiencia humana [8, 9, 10], junto con la filosofía ambiental, que incorpora las dimensiones éticas y ecosistémicas asociadas con el desarrollo tecnológico [11], perspectivas cuyos respectivos énfasis en medición, causalidad, interpretación y responsabilidad abren la posibilidad de examinar cómo estas formas de razonamiento se proyectan sobre las capacidades profesionales requeridas en los sistemas industriales.

Desde esta proyección hacia las capacidades profesionales, las competencias priorizadas en los programas educativos y empresariales adquieren particular relevancia, dado que la formación vinculada con la automatización suele concentrarse en capacidades relacionadas con la operación, integración y optimización de tecnologías, mientras permanecen menos explícitos los supuestos desde los cuales se interpreta un problema y se selecciona la forma de razonamiento utilizada para abordarlo. Una anomalía registrada por un sensor puede requerir verificación cuantitativa y, al mismo tiempo, conducir a la exploración de mecanismos causales no evidentes, a la consideración de las condiciones particulares del entorno de operación o a la valoración de las implicaciones asociadas con una determinada respuesta, relaciones que adquieren nuevas configuraciones conforme los sistemas de inteligencia artificial participan con mayor frecuencia en actividades de supervisión, predicción y apoyo a la toma de decisiones.

Ante esta creciente participación de sistemas inteligentes en los procesos decisionales, las posibilidades de procesamiento y análisis de información se amplían junto con condiciones relacionadas con la opacidad de determinados modelos, la dependencia respecto de los datos empleados y las particularidades del contexto de operación, aspectos que intervienen cuando una respuesta algorítmica cuantitativamente consistente requiere ser interpretada antes de trasladarse a una decisión industrial. Bajo estas condiciones, la verificación empírica puede relacionarse con la exploración causal, mientras la interpretación contextual y la consideración ética aportan otros niveles de lectura sobre una misma situación, configurando una interacción entre formas de razonamiento cuya correspondencia con las capacidades profesionales plantea la

necesidad de examinar cómo los fundamentos filosóficos pueden integrarse en la formación vinculada con la automatización industrial.

Frente a esta necesidad de articulación entre fundamentos filosóficos y capacidades profesionales, el objetivo de la investigación es analizar cómo el positivismo lógico, el realismo crítico, la hermenéutica y la filosofía ambiental interactúan en la formación profesional y proponer un plan integrado de capacidades para los futuros arquitectos tecnológicos de la automatización industrial, considerando la correspondencia entre las perspectivas filosóficas, las formas de razonamiento que promueven y las capacidades asociadas con diferentes niveles de decisión. La propuesta se desarrolla a partir de estas relaciones y busca representar su articulación dentro de una estructura que vincule los fundamentos filosóficos con las capacidades profesionales y las situaciones propias de los sistemas industriales contemporáneos, manteniendo como referencia el carácter documental y conceptual desde el cual se aborda el estudio.

En correspondencia con este propósito, el resto del artículo se organiza de la siguiente manera. La Sección 2 examina los trabajos relacionados con las cuatro líneas de literatura consideradas y los espacios de investigación identificados, mientras que la Sección 3 describe los métodos empleados y precisa el carácter exclusivamente documental de la base empírica del estudio, sin aplicación de instrumentos a personas. La Sección 4 presenta los resultados del análisis documental, la cartografía de la pirámide y la propuesta de modelo integrado con su algoritmo asociado. Posteriormente, la Sección 5 contrasta los hallazgos con la literatura actual y examina las limitaciones vinculadas con el alcance conceptual del estudio, mientras que la Sección 6 presenta las conclusiones y las posibilidades de validación empírica de la propuesta en entornos industriales reales.

2. Trabajos relacionados

En esta sección se analizan los trabajos relacionados con la influencia de la filosofía de la ciencia en la ingeniería y la automatización industrial, considerando aportes vinculados con la formación filosófica en ingeniería, el realismo crítico y la retroducción causal, el diseño centrado en el humano en la Industria 4.0 y 5.0, y la supervisión humana de sistemas inteligentes. La revisión se organiza a partir de estas perspectivas debido a su vinculación con las corrientes filosóficas consideradas en el estudio y con diferentes niveles de análisis que abarcan la formación profesional, la explicación de mecanismos causales, la interacción humano-tecnología y la supervisión de sistemas algorítmicos. A partir de esta organización, los antecedentes se examinan atendiendo tanto a sus aportes como a las relaciones que pueden establecerse entre ellos en torno a la formación y la toma de decisiones en la automatización industrial.

Desde el ámbito de la formación en ingeniería, Lysak [2] desarrolló un análisis teórico de las dimensiones filosóficas presentes en la educación de los ingenieros y relacionó la ausencia de una formación filosófica explícita con dificultades de clasificación ontológica y con puntos ciegos éticos durante el diseño de sistemas técnicos, aportando una base para reconocer que las decisiones profesionales no dependen exclusivamente del conocimiento instrumental. Aunque esta perspectiva permite incorporar la reflexión filosófica en la formación del ingeniero, su desarrollo permanece principalmente en el plano curricular general y no establece una correspondencia entre corrientes filosóficas, capacidades profesionales y niveles específicos de decisión dentro de los sistemas de automatización industrial. Ante este espacio, el presente estudio traslada la discusión hacia el contexto de la automatización mediante la vinculación de los fundamentos filosóficos con capacidades diferenciadas que se sostienen a través de una pirámide de automatización, incorporando además un mecanismo de priorización que busca convertir estos fundamentos en criterios aplicables a situaciones de decisión.

Esta necesidad de trasladar los fundamentos filosóficos hacia formas de razonamiento aplicables encuentra un antecedente relevante en el realismo crítico, particularmente en los trabajos orientados a operacionalizar la explicación causal. Por ejemplo, Rieder et al. [7] propusieron cuatro pasos para codificar la retroducción causal en modelos basados en agentes, convirtiendo un principio filosófico en un procedimiento reproducible aplicado a la simulación computacional de fenómenos sociales, mientras que Mukumbang [6] mostró cómo la retroducción puede integrar métodos cualitativos y cuantitativos mediante la distinción entre eventos observables y mecanismos generativos subyacentes. Estos aportes avanzan en la traducción del razonamiento filosófico hacia procedimientos de análisis, aunque sus aplicaciones no se

orientan específicamente a la toma de decisiones en sistemas de automatización industrial ni establecen su relación con capacidades profesionales diferenciadas. Sobre esta brecha, el presente estudio adapta la lógica retroductiva al análisis de anomalías dentro de procesos automatizados y la incorpora como parte del mecanismo de diagnóstico propuesto, relacionando la identificación de eventos observables con la exploración de posibles mecanismos generativos y con la selección de capacidades requeridas para su interpretación.

Por tanto, la interpretación de estas situaciones no depende únicamente de la identificación de mecanismos causales, sino también de las condiciones en las que interactúan las personas y los sistemas tecnológicos, aspecto que ha adquirido mayor presencia en los estudios sobre Industria 4.0 y 5.0. Nguyen et al. [9] revisaron casos de diseño centrado en el humano en la Industria 4.0 e identificaron una limitada articulación entre estas aproximaciones y los marcos orientados al desarrollo de competencias, mientras que Yaqot et al. [10] analizaron la relación entre industria y participación humana en la Industria 5.0 considerando las tensiones asociadas con la sustitución progresiva del esfuerzo humano por sistemas autónomos. Ambos aportes amplían la comprensión de la automatización al situar la experiencia y la participación humana dentro del sistema productivo, aunque la relación entre estas dimensiones y los fundamentos hermenéuticos que pueden orientar las capacidades de interpretación profesional permanece menos desarrollada. De esta manera, el presente trabajo aborda esta brecha mediante la incorporación de la perspectiva hermenéutica en el plan de capacidades, vinculándola con situaciones en las que el profesional requiere interpretar el contexto de operación, las interacciones humano-tecnología y las implicaciones asociadas con decisiones que trascienden el análisis exclusivamente cuantitativo.

A ello, se suma la creciente participación humana en tareas de interpretación que adquiere otra dimensión cuando los sistemas inteligentes asumen funciones de predicción, recomendación o decisión y el profesional pasa de intervenir directamente en el proceso a supervisar sus resultados. En este contexto, Laux [3] analizó los requisitos de supervisión humana asociados con el Reglamento de Inteligencia Artificial de la Unión Europea, mientras que Sterz et al. [12] y Langer et al. [4] desarrollaron marcos interdisciplinarios para examinar las condiciones que determinan la efectividad de dicha supervisión frente a sistemas caracterizados por diferentes grados de opacidad, a lo que Vallor y Vierkant [13] incorporaron el análisis de la vulnerabilidad relacional y de los vacíos de responsabilidad que pueden surgir entre las personas y los sistemas socio-técnicos. En efecto, estos trabajos permiten desplazar la atención desde la simple presencia de un operador hacia las condiciones necesarias para ejercer una supervisión significativa, aunque prestan menor atención a la manera en que la formación filosófica del profesional puede contribuir a interpretar, cuestionar y valorar las respuestas generadas por sistemas inteligentes. Frente a este espacio, el manuscrito incorpora la supervisión humana dentro del mecanismo de escalamiento del modelo propuesto y la relaciona con capacidades de razonamiento que combinan verificación, explicación causal, interpretación contextual y valoración de las implicaciones asociadas con la decisión.

De esta manera, la convergencia de estos trabajos previos permite observar que la literatura ha desarrollado aportes relevantes para comprender, de manera diferenciada, la formación filosófica en ingeniería [2], la operacionalización del razonamiento causal [7, 6], la participación humana en los sistemas industriales [9, 10] y las condiciones asociadas con la supervisión y la responsabilidad frente a sistemas inteligentes [3, 12, 4, 13]. Sin embargo, estos desarrollos permanecen distribuidos entre marcos conceptuales y campos de aplicación diferentes, dejando menos explorada su articulación dentro de una estructura que relacione corrientes filosóficas, formas de razonamiento, capacidades profesionales y niveles de decisión propios de la automatización industrial. Por ello, el estudio atiende esta brecha mediante una propuesta integrada de capacidades que articula las perspectivas filosóficas con las formas de razonamiento requeridas en la automatización industrial, estableciendo su correspondencia con diferentes niveles de decisión y trasladando esta articulación hacia una estructura que incorpora mecanismos de priorización y diagnóstico. De esta forma, la propuesta busca superar el tratamiento fragmentado de estas dimensiones mediante un marco que permita relacionar fundamentos filosóficos, capacidades profesionales y situaciones de decisión dentro de los sistemas industriales contemporáneos.

3. Metodología

En esta sección se describe el procedimiento seguido para identificar, seleccionar, analizar e integrar la literatura que sustenta la construcción del marco de capacidades propuesto para la toma de decisiones en sistemas inteligentes de automatización industrial. Dado que el objetivo del estudio no consiste únicamente en caracterizar corrientes filosóficas, sino en establecer correspondencias fundamentadas entre sus principios, las capacidades profesionales y los distintos niveles de decisión presentes en la Industria 4.0, se definió una estrategia metodológica que permitiera avanzar de manera trazable desde la evidencia documental hasta la formulación del marco conceptual. Para ello, la investigación se desarrolló desde un enfoque epistemológico transdisciplinario y mediante una metodología filosófico-analítica de carácter documental, en la que las unidades de análisis estuvieron constituidas exclusivamente por publicaciones académicas y documentos especializados.

Sobre esta base, se adoptó un diseño de revisión documental de alcance [14], adecuado para examinar aportes provenientes de distintos campos de conocimiento, establecer relaciones entre sus fundamentos y sintetizarlos dentro de una estructura conceptual integradora. El procedimiento metodológico se estructuró mediante cinco componentes articulados: búsqueda documental y construcción del corpus, análisis documental dirigido y codificación temática, construcción del modelo analógico, aplicación del marco a la toma de decisiones en automatización industrial e incorporación de la filosofía ambiental. Para ello, la Figura 1 representa la articulación del procedimiento metodológico. El recorrido parte de la identificación y conformación del corpus, continúa con su codificación y análisis conceptual y, posteriormente, traslada las categorías obtenidas hacia la estructura de automatización y los niveles de toma de decisiones. La filosofía ambiental se incorpora para ampliar el análisis de las capacidades más allá de la eficiencia técnica y considerar las implicaciones ambientales de las decisiones industriales. Aunque el procedimiento establece una secuencia analítica para asegurar su trazabilidad, la construcción conceptual conserva un carácter iterativo, de modo que las correspondencias identificadas durante la integración del modelo permiten revisar y precisar las categorías obtenidas durante el análisis documental.

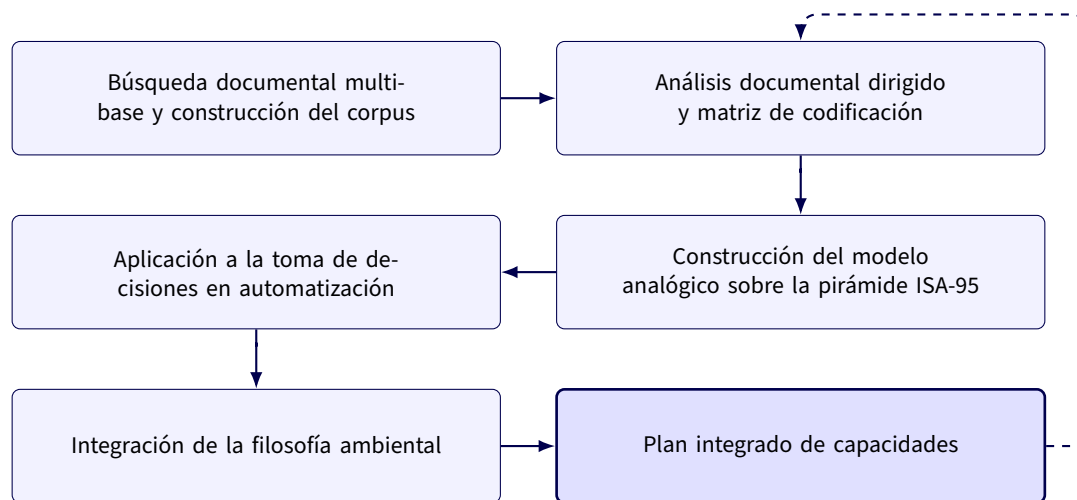


Figura 1. Procedimiento metodológico seguido para la construcción del marco integrado de capacidades

3.1 Búsqueda documental y construcción del corpus

La primera parte del procedimiento estuvo orientada a conformar un corpus documental que proporcionara el sustento conceptual necesario para relacionar las corrientes filosóficas estudiadas con las capacidades requeridas en la toma de decisiones dentro de sistemas inteligentes de automatización industrial. Para ello, se desarrolló una búsqueda estructurada en Scopus, Web of Science y Google Scholar, combinando términos relacionados con positivismo, realismo crítico, hermenéutica, filosofía ambiental, automatización industrial e Industria 4.0. La consulta en diferentes fuentes permitió ampliar la cobertura temática y recuperar contribuciones procedentes tanto del ámbito filosófico como de la ingeniería y la automatización. La estrategia siguió los principios de búsqueda en múltiples bases de datos utilizados en revisiones de alcance [14], junto con recomendaciones dirigidas a mantener la transparencia en la identificación y documentación de las fuentes [15].

La búsqueda se delimitó al periodo 2020–2026 con el propósito de considerar desarrollos recientes vinculados con la transformación de los sistemas industriales. Como criterios de inclusión se establecieron la correspondencia con el periodo definido, la disponibilidad del texto completo y la vinculación sustantiva con al menos uno de los ejes conceptuales del estudio: corrientes filosóficas, capacidades profesionales, automatización industrial, sistemas inteligentes o toma de decisiones. Posteriormente, los documentos recuperados fueron sometidos a un proceso de depuración y revisión de pertinencia, mediante el cual se excluyeron registros duplicados, publicaciones sin relación directa con el problema de investigación y trabajos que no aportaban fundamentos conceptuales o metodológicos para la construcción del marco. En la selección final se priorizaron aquellos estudios que, además de abordar los ejes definidos, permitieran establecer relaciones conceptuales entre las perspectivas filosóficas y su aplicación al contexto de la automatización industrial.

De este modo, la construcción del corpus siguió una secuencia de identificación, depuración, evaluación de pertinencia y selección, como se representa en la Figura 2. La aplicación de estos criterios condujo a un corpus definitivo de 22 publicaciones, que constituyeron las unidades documentales empleadas en el análisis dirigido y la codificación temática. Esta delimitación permitió establecer una base documental trazable para identificar las contribuciones de cada corriente filosófica, contrastarlas con los requerimientos de los sistemas de automatización y sustentar posteriormente las correspondencias incorporadas al marco de capacidades.

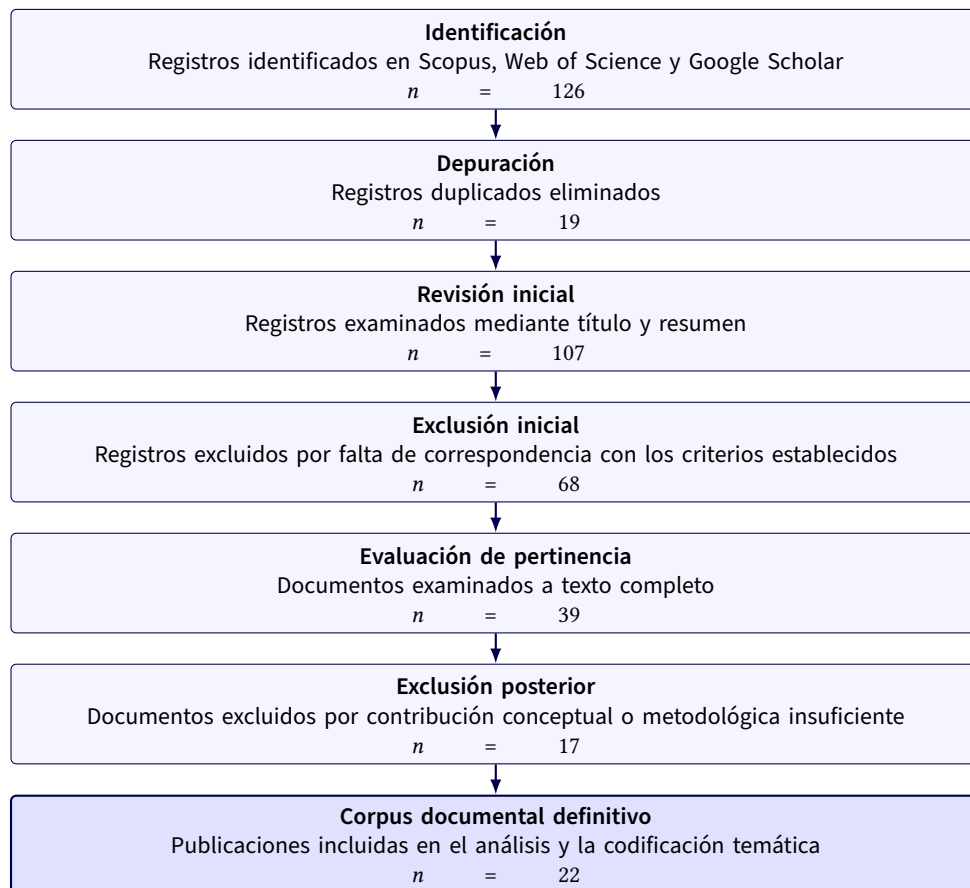


Figura 2. Proceso de identificación, depuración, evaluación y selección del corpus documental.

3.2 Análisis documental dirigido y matriz de codificación

Una vez constituido el corpus, se desarrolló un análisis documental dirigido con el propósito de transformar la información contenida en las fuentes en categorías comparables y susceptibles de integración. Cada documento fue examinado atendiendo a su contribución conceptual y clasificado según la corriente filosófica predominante que sustentaba su argumentación: positivismo, realismo crítico, hermenéutica o filosofía ambiental. De manera complementaria, las publicaciones cuya principal contribución correspondía al dominio tecnológico o metodológico fueron codificadas según su función de soporte, particularmente en relación

con capacidades dinámicas, arquitectura de automatización y supervisión de sistemas de inteligencia artificial.

Este procedimiento permitió establecer una estructura común para comparar documentos provenientes de campos disciplinares diferentes y reconocer convergencias entre fundamentos filosóficos, capacidades y necesidades de decisión. La estrategia es coherente con enfoques recientes que articulan el análisis cualitativo dirigido con procedimientos sistemáticos de síntesis documental [16]. Como resultado se construyó la matriz de codificación temática presentada en la Tabla 2, utilizada posteriormente como mecanismo de trazabilidad para vincular las fuentes del corpus con las categorías que sustentan la construcción del modelo.

3.3 Construcción del modelo analógico

A partir de las categorías obtenidas durante el análisis documental, el siguiente componente metodológico consistió en establecer correspondencias entre los fundamentos epistemológicos identificados y las capacidades requeridas en los sistemas contemporáneos de automatización industrial. Esta operación permitió trasladar conceptos inicialmente formulados en el plano filosófico hacia situaciones concretas de interpretación, diagnóstico y decisión tecnológica, sin asumir una equivalencia directa entre ambos dominios, sino relaciones conceptuales sustentadas en las funciones que cada perspectiva aporta al proceso decisional. Para organizar estas correspondencias se utilizó la pirámide de automatización ISA-95 como estructura de referencia, debido a que permite diferenciar niveles de operación, supervisión, gestión y decisión dentro de los sistemas industriales. Su pertinencia ha sido examinada en arquitecturas recientes de referencia para la Industria 4.0 [1] y en aplicaciones vinculadas con la construcción de gemelos digitales en sistemas didácticos de manufactura [17]. Sobre esta estructura se cartografiaron las relaciones entre niveles de automatización, necesidades de interpretación y capacidades asociadas a las distintas corrientes filosóficas, cuya representación se presenta en la Figura 3.

Por otro lado, la consistencia conceptual de esta articulación se contrastó posteriormente con la teoría de capacidades dinámicas revisada por Cristofaro *et al.* [18], particularmente con su planteamiento sobre la necesidad de integrar, construir y reconfigurar competencias internas y externas frente a entornos cambiantes. En efecto, esta perspectiva proporciona un fundamento organizacional para interpretar las capacidades propuestas no como conocimientos estáticos, sino como recursos susceptibles de adaptación ante la evolución tecnológica característica de la Industria 4.0.

3.4 Aplicación a la toma de decisiones en automatización industrial

Una vez establecidas las correspondencias conceptuales, se examinó su aplicabilidad en situaciones de toma de decisiones propias de los diferentes niveles de automatización industrial. Para ello, las decisiones fueron diferenciadas según su finalidad y contexto de actuación, distinguiendo aquellas de naturaleza reactiva, vinculadas con la respuesta ante eventos, fallas o desviaciones del proceso, de aquellas de carácter progresivo, asociadas con la mejora continua, la adaptación y la innovación. Esta distinción se encuentra en consonancia con desarrollos recientes sobre sistemas de automatización y supervisión para la Industria 4.0 [19].

De este modo, el análisis permitió relacionar cada nivel de la pirámide con el tipo de decisión predominante y con las perspectivas filosóficas que ofrecen mayor capacidad explicativa o interpretativa ante cada situación. Estas relaciones fueron sistematizadas en la Tabla 3 y sirvieron posteriormente como fundamento para construir el algoritmo de diagnóstico filosófico-causal presentado en la Sección 4.5.1. Su lógica establece un recorrido progresivo que comienza con la explicación cuantitativa del comportamiento observado, continúa con la identificación de mecanismos causales y la interpretación contextual y, cuando la naturaleza o las consecuencias de la decisión lo requieren, incorpora la supervisión humana. De igual manera, la Figura 5 sintetiza este procedimiento y muestra cómo las perspectivas filosóficas se articulan como recursos complementarios para ampliar la capacidad de diagnóstico y decisión.

3.5 Filosofía ambiental en la automatización industrial

El último componente metodológico incorporó la filosofía ambiental como criterio para analizar las implicaciones de las decisiones tecnológicas dentro de los sistemas de automatización industrial. Para ello, las categorías obtenidas del análisis documental se examinaron considerando no solo su contribución al desempeño técnico y operativo, sino también su relación con el uso responsable de recursos, la eficiencia energética, la sostenibilidad de los procesos y las consecuencias ambientales derivadas de las decisiones industriales. Este procedimiento permitió ampliar la concepción de las capacidades requeridas en la Industria 4.0, al reconocer que la selección, configuración y operación de tecnologías inteligentes implican responsabilidades que trascienden los resultados inmediatos y alcanzan sus efectos ecológicos futuros, en consonancia con planteamientos de ética ambiental aplicada a la ingeniería [11]. A partir de esta integración, la responsabilidad ambiental se estableció como cuarto pilar del marco propuesto, complementando la explicación cuantitativa, la comprensión causal y la interpretación contextual desarrolladas en los componentes anteriores.

4. Resultados

Esta sección presenta los resultados derivados del análisis documental y de la integración conceptual desarrollada en el estudio. La exposición sigue una secuencia que permite avanzar desde la identificación de las capacidades asociadas con cada corriente filosófica hasta su aplicación en los distintos niveles de automatización industrial. En primer lugar, se examinan las contribuciones del positivismo, el realismo crítico y la hermenéutica; posteriormente, estas relaciones se sintetizan mediante matrices de correspondencia y codificación temática. Sobre esta base se construye la cartografía filosófica de la pirámide de automatización ISA-95 y se formula el plan integrado de capacidades. Por último, se presenta el algoritmo de diagnóstico filosófico-causal, junto con la incorporación de la filosofía ambiental y la supervisión humana como dimensiones complementarias para la toma de decisiones en sistemas inteligentes.

4.1 Influencia del positivismo lógico en competencias operativas

El análisis del corpus permitió identificar que una parte importante de las capacidades operativas descritas en la literatura sobre automatización se sustenta en principios vinculados con la observación, medición y verificación cuantitativa. Indicadores como la eficacia global de los equipos, las tasas de defecto o el retorno de la inversión representan esta orientación, debido a que proporcionan criterios objetivos para evaluar el comportamiento de los procesos y establecer acciones correctivas. Desde esta perspectiva, el positivismo aporta una base necesaria para garantizar control, predictibilidad y consistencia en los niveles operativos de la automatización. Sin embargo, su aplicación aislada puede resultar insuficiente cuando la explicación de una anomalía requiere comprender estructuras causales que no son directamente observables. Por tanto, esta limitación coincide con lo señalado por Lysak [2], quien advierte que la ausencia de fundamentos filosóficos explícitos en la formación en ingeniería puede conducir a interpretaciones ontológicas insuficientes sobre las capacidades reales de un sistema, particularmente cuando intervienen componentes de aprendizaje automático cuya lógica interna no puede explicarse solamente mediante resultados cuantificables. Por tanto, el positivismo constituye una base fundamental para las competencias operativas, pero requiere ser complementado con formas de razonamiento capaces de abordar las causas estructurales, los contextos de uso y las implicaciones de las decisiones tecnológicas.

De igual manera, esta necesidad de complementariedad también se relaciona con lo planteado por Ferdman y Ratti [5], quienes muestran que el ideal de neutralidad valorativa heredado del positivismo puede dificultar que los profesionales reconozcan las dimensiones éticas y sociales incorporadas en los propios diseños tecnológicos. Por ello, el plan de capacidades propuesto conserva las competencias cuantitativas y técnicas como base del desempeño industrial, pero evita considerarlas suficientes para responder a situaciones que demandan explicaciones causales, interpretación contextual o evaluación de consecuencias.

4.2 Aportación del realismo crítico a la resiliencia estructural

En este apartado se analiza la aportación del realismo crítico a la comprensión de problemas que no pueden resolverse observando únicamente los datos visibles de un sistema industrial. Mientras el positivismo permite medir una desviación, identificar una alarma o comprobar que una variable superó determinado um-

bral, el realismo crítico dirige el análisis hacia una pregunta adicional: ¿qué mecanismos o condiciones hicieron que ese evento ocurriera? Esta perspectiva resulta especialmente relevante en sistemas de automatización interconectados, donde una misma falla puede originarse por diferentes combinaciones de factores. Para abordar esta complejidad se recurre a la retroducción causal, entendida como un razonamiento que parte del evento observado y busca identificar los mecanismos capaces de explicarlo [6]. Así, ante una detención recurrente de una máquina, el profesional no solo registra su frecuencia, sino que examina posibles relaciones entre sensores, programación, mantenimiento y condiciones de operación.

Esta manera de analizar los problemas contribuye a la resiliencia estructural porque permite pasar de corregir los efectos inmediatos de una anomalía a comprender y atender las causas que favorecen su aparición. Rieder et al. [7] muestran que la identificación de estos mecanismos generativos puede organizarse mediante procedimientos explícitos, haciendo posible trasladar este fundamento filosófico hacia procesos sistemáticos de diagnóstico. De esta manera, en el marco propuesto, el realismo crítico no sustituye al positivismo, sino que amplía su capacidad explicativa cuando los indicadores cuantitativos dejan preguntas sin resolver, permitiendo formular posibles causas, contrastarlas con la evidencia disponible y orientar acciones sobre la estructura que origina el problema. Esta complementariedad constituye uno de los fundamentos del Algoritmo 1, desarrollado posteriormente en la Sección 4.5.1, donde este razonamiento se incorpora como parte del proceso de diagnóstico y toma de decisiones.

4.3 Impacto de la hermenéutica en el diseño centrado en el humano

Después de ampliar el análisis desde la medición de los eventos hacia la comprensión de los mecanismos que los producen, la hermenéutica incorpora una tercera perspectiva que permite considerar las condiciones humanas y contextuales presentes en el funcionamiento de los sistemas automatizados. Desde esta mirada, una decisión no depende únicamente de los datos disponibles ni de la identificación de sus posibles causas, sino también de la manera en que los operadores interpretan la información, comprenden una situación y actúan frente a ella; por ello, esta capacidad adquiere especial relevancia cuando los sistemas inteligentes participan en tareas de supervisión, recomendación y apoyo a la toma de decisiones, donde la interacción entre personas y tecnología se vuelve parte del propio proceso decisional. Esta orientación coincide con estudios recientes sobre diseño centrado en el humano en escenarios de Industria 4.0 y 5.0, que destacan la necesidad de superar una concepción exclusivamente instrumental de la automatización [9, 10].

En este sentido, la hermenéutica aporta conceptos como la precomprensión y la interpretación del contexto, mediante los cuales una falla o una respuesta inesperada deja de entenderse exclusivamente como un acontecimiento técnico y comienza a analizarse también desde la experiencia de quienes interactúan con el sistema. Grunwald et al. [8] sostienen que las situaciones de ruptura permiten reconsiderar la relación entre las personas, los sistemas y las circunstancias en las que ambos interactúan; trasladado a la automatización industrial, esto significa que una alarma incomprendida, una interfaz poco utilizada o la resistencia de un operador frente a una recomendación algorítmica pueden revelar dificultades que no necesariamente se originan en la configuración técnica. De esta manera, la hermenéutica completa el recorrido iniciado con la medición positivista y ampliado mediante la explicación causal del realismo crítico, incorporando la comprensión del contexto como una capacidad necesaria para interpretar situaciones de mayor complejidad.

Esta articulación permite comprender las tres corrientes como perspectivas complementarias dentro de un mismo proceso de toma de decisiones, pues mientras el positivismo aporta criterios para observar, medir y verificar el comportamiento del sistema, el realismo crítico permite profundizar en los mecanismos que pueden explicar dicho comportamiento y la hermenéutica incorpora las condiciones humanas y contextuales que intervienen en su interpretación. Con el propósito de hacer explícita esta complementariedad, la Tabla 1 organiza las principales correspondencias identificadas entre cada corriente filosófica, su enfoque de análisis, la competencia que prioriza y su incidencia sobre las decisiones en automatización industrial.

Tabla 1. Matriz de corrientes filosóficas y su impacto en las capacidades de automatización

Corriente filosófica	Enfoque principal	Competencia priorizada	Impacto en la toma de decisiones	Ref.
Positivismo	Empirismo, cuantificación, leyes lógicas	Técnicas duras, programación, control exacto	Optimización lineal y eficiencia operativa	[5]
Realismo crítico	Ontología estratificada, retroducción causal	Análisis de fallos, resiliencia sistémica	Diagnóstico estructural profundo ante anomalías	[6, 7]
Hermenéutica	Interpretación del significado y los contextos	Diseño centrado en el humano, innovación	Integración socio-técnica y adaptabilidad cultural	[8, 9, 10]

Como se aprecia en la Tabla 1, la diferencia entre las corrientes no radica en que una sustituya a las demás, sino en el tipo de problema que cada una permite abordar con mayor profundidad. Así, la medición y el control proporcionados por el positivismo constituyen una base para reconocer el comportamiento observable del sistema, sobre la cual el realismo crítico amplía la explicación hacia los mecanismos causales, mientras que la hermenéutica incorpora la interpretación de las condiciones humanas y contextuales. En consecuencia, el aumento de la complejidad de una decisión requiere ampliar progresivamente las capacidades movilizadas, lo que permite pasar de una respuesta predominantemente técnica hacia una comprensión que articula datos, causas y contexto.

Ahora bien, establecer esta correspondencia conceptual requiere también mostrar de qué manera fue sustentada por el corpus documental, ya que la construcción del marco no se apoyó exclusivamente en las publicaciones directamente vinculadas con positivismo, realismo crítico y hermenéutica. Por el contrario, durante la codificación se identificaron contribuciones complementarias relacionadas con capacidades dinámicas, arquitectura de automatización, supervisión humana, filosofía ambiental y fundamentos metodológicos, las cuales permitieron conectar las corrientes filosóficas con problemas propios de los sistemas inteligentes industriales. Esta organización se presenta en la Tabla 2, donde se muestra la distribución temática de las 22 fuentes que conformaron el corpus documental.

Tabla 2. Matriz de codificación temática del corpus documental

Categoría temática	Núm. de fuentes	Referencias
Positivismo y formación técnica	1	[5]
Realismo crítico y retroducción	2	[6, 7]
Hermenéutica y diseño centrado en el humano	3	[8, 9, 10]
Filosofía ambiental y ética de la ingeniería	1	[11]
Capacidades dinámicas organizacionales	2	[18, 20]
Supervisión humana y agencia de sistemas de IA	6	[3, 12, 4, 13, 21, 22]
Arquitectura de automatización industrial	3	[1, 17, 19]
Filosofía y formación en ingeniería	1	[2]
Método de revisión documental	3	[16, 15, 14]

A partir de la distribución presentada en la Tabla 2, se observa que el sustento documental del marco se extiende más allá de las tres corrientes filosóficas analizadas de manera directa, debido a que su aplicación al contexto industrial exige establecer vínculos con otros campos que explican cómo esas formas de razonamiento pueden trasladarse hacia capacidades concretas. En particular, la mayor presencia de estudios relacionados con supervisión humana y agencia de sistemas de IA evidencia la importancia que adquiere el juicio profesional cuando los sistemas inteligentes incrementan su participación en la toma de decisiones; al mismo tiempo, las fuentes sobre arquitectura de automatización proporcionan la estructura tecnológica necesaria para situar dichas capacidades en niveles concretos de operación, supervisión, planificación y gestión. Esta conexión permite avanzar hacia la cartografía filosófica de la pirámide de automatización desarrollada en la siguiente subsección.

4.4 Cartografía filosófica de la pirámide de automatización

Una vez identificadas las contribuciones particulares de cada corriente, el siguiente resultado consistió en determinar cómo estas formas de razonamiento pueden relacionarse con los diferentes niveles de la automatización industrial. Para ello, se utilizó como referencia la organización jerárquica de ISA-95 [1, 17], estableciendo una correspondencia entre nivel tecnológico, naturaleza predominante de la decisión y perspectiva filosófica requerida. La Tabla 3 presenta esta cartografía y permite observar cómo la necesidad de integrar diferentes formas de razonamiento aumenta a medida que las decisiones trascienden la respuesta operativa inmediata.

Tabla 3. Niveles de la pirámide de automatización, tipo de decisión y corriente filosófica dominante

Nivel	Tipo de decisión	Corriente dominante	Ref.
Campo (sensores y actuadores)	Reactiva	Positivismo	[1, 17]
Control (PLC/DCS)	Programada	Positivismo con anticipación causal	[1, 17]
Supervisión (SCADA/HMI)	Táctica	Realismo crítico + hermenéutica	[19, 8]
Planificación (MES/MOM)	Estratégica	Realismo crítico + hermenéutica	[19, 10]
Gestión (ERP)	Corporativa	Integración de las tres corrientes + filosofía ambiental	[1, 11]

La Tabla 3 evidencia una progresión en la complejidad del razonamiento requerido. En los niveles de campo y control predominan las decisiones sustentadas en mediciones, reglas y condiciones previamente establecidas, mientras que en los niveles de supervisión y planificación aumenta la necesidad de explicar mecanismos causales e interpretar el contexto de interacción con los operadores. En el nivel de gestión, donde las decisiones afectan al conjunto de la organización y pueden producir consecuencias de mayor alcance, la propuesta converge hacia una integración de las tres corrientes junto con consideraciones ambientales. Esta progresión se representa espacialmente en la Figura 3, donde la estructura ISA-95 permite visualizar la transición desde una base predominantemente positivista hacia niveles en los que se requiere una combinación progresiva de perspectivas filosóficas.

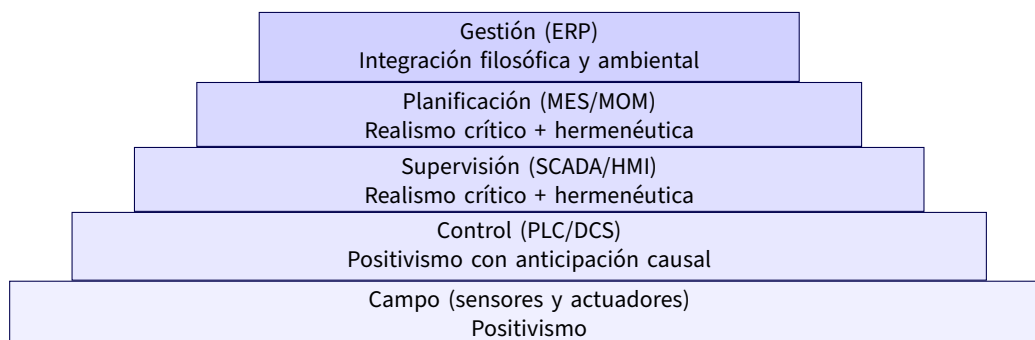


Figura 3. Pirámide de automatización ISA-95 cartografiada según la corriente filosófica dominante en cada nivel de decisión [1]

En efecto, la Figura 3 permite observar que las corrientes filosóficas no se asignan como compartimentos independientes. Por el contrario, su participación se amplía conforme aumenta el alcance de la decisión. El positivismo conserva su función en todos los niveles debido a que la información cuantitativa continúa siendo necesaria; sin embargo, en niveles superiores resulta insuficiente sin una comprensión causal e interpretativa capaz de incorporar condiciones organizacionales y humanas. En esta dirección y con el propósito de hacer visible esta superposición, la Figura 4 representa conceptualmente el grado de convergencia entre el razonamiento cuantitativo y el contextual a través de los niveles de automatización. El color azul identifica el predominio de la verificación empírica y cuantitativa, mientras que el rojo representa la creciente necesidad de interpretación contextual. La zona de intersección muestra el espacio en el que ambos enfoques requieren ser articulados mediante razonamiento causal.

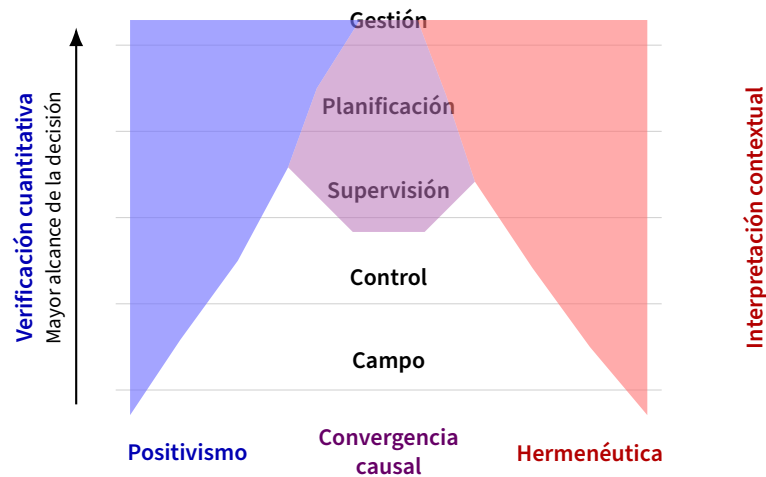


Figura 4. Convergencia conceptual de las perspectivas filosóficas según el nivel de automatización industrial

De esta manera, la Figura 4 refuerza una idea central del marco propuesto: el ascenso dentro de la estructura de automatización no supone abandonar las capacidades cuantitativas desarrolladas en los niveles inferiores, sino incorporar progresivamente otras formas de razonamiento. En este sentido, la zona de convergencia simboliza la función articuladora del realismo crítico, que permite trasladar la observación empírica hacia la explicación de mecanismos y, posteriormente, hacia la interpretación contextual requerida por decisiones de mayor alcance.

4.5 Propuesta del plan de capacidades integrado

Una vez establecida la relación entre las corrientes filosóficas y los distintos niveles de la pirámide de automatización, el siguiente paso consiste en traducir esas correspondencias en capacidades concretas que puedan orientar la formación y el desempeño profesional. Bajo esta lógica, el plan propuesto no considera las competencias como conocimientos técnicos aislados, sino como capacidades que se complementan conforme aumenta la complejidad del problema que debe resolverse. Así, una situación operativa puede requerir inicialmente medición y control, mientras que una falla recurrente exige profundizar en sus posibles causas y una decisión que involucra a personas demanda, además, comprender el contexto en el que ocurre. A estas capacidades se incorpora la responsabilidad ambiental como criterio transversal, debido a que las decisiones tecnológicas también pueden generar consecuencias sobre los recursos y el entorno. Esta organización se sintetiza en la Tabla 4, donde cada capa se relaciona con una corriente filosófica, un conjunto de competencias y un ejemplo de aplicación industrial.

Tabla 4. Plan de capacidades integrado por capa, corriente filosófica y ejemplo de aplicación

Capa	Corriente	Competencias clave	Ejemplo de aplicación	Ref.
Operativa	Positivismo	Programación PLC/SCADA, calibración de sensores, control estadístico de procesos	Ajuste de parámetros ante una desviación de OEE	[5]
Táctica	Realismo crítico	Retroducción causal, análisis de modos de fallo y efectos (AMFE)	Diagnóstico de un mecanismo generativo tras fallos recurrentes	[7]
Estratégica	Hermenéutica	Interpretación contextual, diseño de interfaces centradas en el humano	Rediseño de alertas SCADA tras resistencia del operador	[9, 10]
Transversal (ético)	Filosofía ambiental	Evaluación de huella ecológica, justicia intergeneracional	Priorización de inversiones según impacto ambiental	[11]

Como se observa en la Tabla 4, el plan sigue una progresión en la que cada capa amplía las capacidades de la anterior en lugar de sustituirlas. De este modo, la capa operativa proporciona las herramientas para programar, medir y controlar el proceso, mientras que la capa táctica incorpora la retroducción causal para investigar qué mecanismos pueden estar detrás de una falla recurrente [7]; cuando la decisión alcanza situaciones en las que intervienen operadores, interfaces y condiciones particulares de uso, la capa estratégica añade la interpretación contextual y el diseño centrado en el humano [9, 10]. A su vez, la filosofía ambiental atraviesa las tres capas para incorporar una pregunta adicional a la decisión: además de determinar si una solución funciona y responde al problema identificado, es necesario considerar qué consecuencias puede producir sobre los recursos, el ambiente y las generaciones futuras [11].

Por tanto, esta integración implica que las capacidades requeridas no permanecen fijas, ya que los problemas industriales, las tecnologías disponibles y las condiciones de operación cambian con el tiempo; por esta razón, la estructura propuesta se relaciona con la teoría de capacidades dinámicas [18], según la cual las competencias deben poder combinarse, adaptarse y reconfigurarse frente a nuevas circunstancias. En la misma dirección, Teece [20] vincula la innovación tecnológica con la capacidad de las organizaciones para transformar y articular sus recursos y competencias, lo que permite comprender el plan no como una secuencia rígida de conocimientos, sino como una estructura adaptable. Sobre esta base, el siguiente paso consiste en convertir dicha integración conceptual en un procedimiento que oriente qué tipo de razonamiento debe emplearse ante una anomalía concreta, propósito que da lugar al algoritmo de diagnóstico filosófico-causal presentado a continuación.

4.5.1 Algoritmo de diagnóstico filosófico-causal

Para trasladar el plan de capacidades desde su formulación conceptual hacia una secuencia de razonamiento aplicable ante anomalías industriales, se desarrolló el Algoritmo 1. Su finalidad es determinar qué tipo de análisis debe priorizarse conforme aumenta la dificultad para explicar un evento mediante los mecanismos habituales de control. El procedimiento comienza con la verificación cuantitativa, continúa con la identificación de mecanismos causales cuando esa verificación resulta insuficiente, incorpora la interpretación del contexto humano-máquina cuando corresponde y, finalmente, establece una instancia de supervisión humana para decisiones mediadas por sistemas de inteligencia artificial cuya actuación no resulte suficientemente transparente.

Algorithm 1 Diagnóstico filosófico-causal para la toma de decisiones ante una anomalía en automatización industrial

Require: Evento empírico e detectado en el nivel L de la pirámide de automatización

Ensure: Decisión d y corriente filosófica que la respalda

- 1: **Verificación positivista:** evaluar si e se explica mediante umbrales cuantitativos establecidos (OEE, tasas de defecto, límites de control)
 - 2: **if** e queda explicado por umbrales cuantitativos **then**
 - 3: **return** $d \leftarrow$ acción correctiva estándar (positivismo)
 - 4: **else**
 - 5: **Retroducción causal:** descomponer e en agentes, eventos y condiciones [7]
 - 6: Conjeturar mecanismos generativos candidatos y eliminar aquellos que no resulten plausibles mediante contraste con la evidencia histórica del proceso
 - 7: **if** se identifica un mecanismo generativo único y plausible **then**
 - 8: **return** $d \leftarrow$ acción estructural dirigida al mecanismo (realismo crítico)
 - 9: **else**
 - 10: **Interpretación hermenéutica:**
 - 11: **if** e involucra interacción humano-máquina (HMI/SCADA) **then**
 - 12: Interpretar el contexto operativo, la carga cognitiva del operador y el significado de la alerta [8]
 - 13: **return** $d \leftarrow$ rediseño de interfaz o procedimiento (hermenéutica)
 - 14: **end if**
 - 15: **end if**
 - 16: **end if**
 - 17: **if** e involucra un componente de inteligencia artificial de decisión opaca **then**
 - 18: **Escalación por supervisión humana:** suspender la ejecución autónoma y transferir la decisión a un operador con formación filosófica integrada [3, 4, 12]
 - 19: **end if**
 - 20: **return** d , corriente(s) filosófica(s) que la fundamentan y registro trazable de la decisión
-

La secuencia definida en el Algoritmo 1 responde a una lógica de escalamiento progresivo. Si el evento puede ser comprendido mediante indicadores y umbrales establecidos, la decisión permanece en el plano operativo; cuando estos elementos resultan insuficientes, se activa la búsqueda de mecanismos causales y, si la situación involucra factores de interacción humano-máquina, el diagnóstico se amplía hacia la interpretación contextual. La supervisión humana constituye el nivel de resguardo cuando intervienen sistemas inteligentes cuya recomendación no puede justificarse de forma suficientemente clara. Este principio se apoya en los trabajos sobre supervisión humana efectiva de Langer et al. [4] y Sterz et al. [12], así como en los criterios de gobernanza algorítmica planteados por Laux [3].

Para facilitar la interpretación de esta secuencia, la Figura 5 traslada la lógica del Algoritmo 1 a una representación visual que muestra el recorrido desde la detección del evento hasta el registro de la decisión y de su fundamento filosófico.

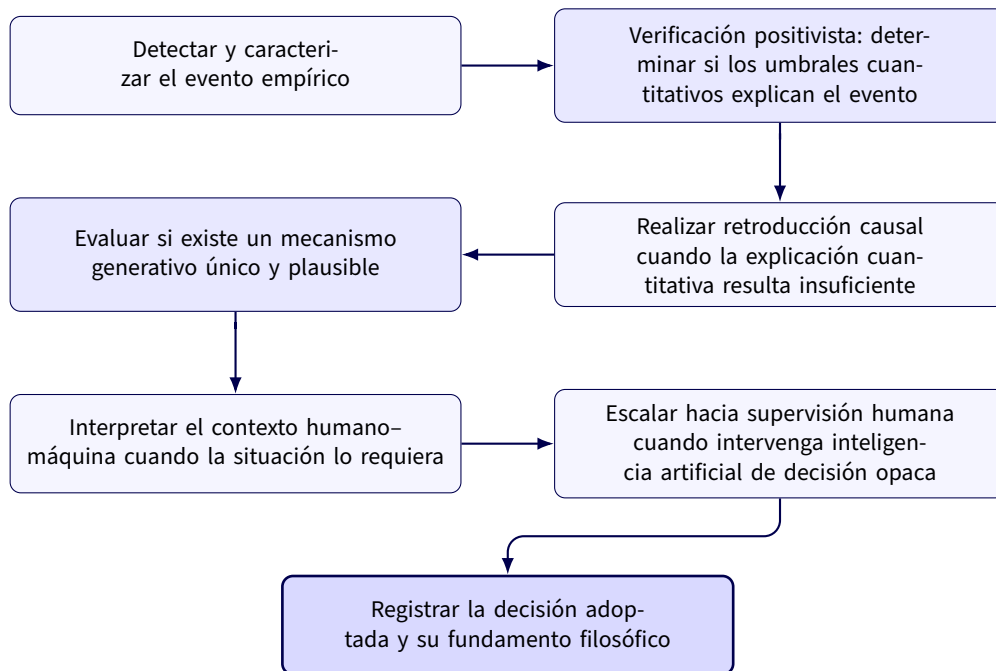


Figura 5. Representación del procedimiento de diagnóstico filosófico-causal definido en el Algoritmo 1

Como muestra la Figura 5, el procedimiento no exige aplicar simultáneamente todas las perspectivas filosóficas, sino incorporar niveles adicionales de razonamiento cuando la naturaleza del problema supera la capacidad explicativa del nivel anterior. En términos prácticos, esta organización responde a la pregunta sobre qué debe atender primero un equipo de automatización ante una anomalía y cuándo debe ampliar el diagnóstico. De esta manera, la propuesta evita tanto reducir todo problema a indicadores cuantitativos como trasladar de forma automática la responsabilidad hacia un sistema inteligente. De esta manera, el Algoritmo 1 constituye una propuesta conceptual derivada de la integración teórica realizada en el estudio y, por tanto, no pretende sustituir los protocolos específicos de seguridad, mantenimiento o control existentes en una instalación industrial. Su contribución consiste en proporcionar una secuencia explícita y trazable para identificar el tipo de razonamiento requerido y documentar el fundamento de la decisión adoptada. La evaluación empírica de su comportamiento en escenarios reales de planta se plantea como una línea posterior de investigación.

4.6 Filosofía ambiental con ética y valores en la automatización industrial

Una vez integradas las capacidades para medir los eventos, comprender sus causas e interpretar el contexto en el que ocurren, resulta necesario ampliar el análisis hacia las consecuencias que una decisión tecnológica puede producir más allá del funcionamiento inmediato del sistema. Desde esta perspectiva, una solución de automatización no debería valorarse solamente por su capacidad para aumentar la productividad, reducir errores o mejorar el control, sino también por la forma en que utiliza los recursos y por los efectos que puede generar sobre el entorno. En este punto, la ética ambiental aplicada a la ingeniería [11] aporta criterios para considerar el consumo energético, la eficiencia en el uso de materiales, el impacto ecológico y la

responsabilidad frente a las generaciones futuras; por ello, la filosofía ambiental se incorpora al marco como una dimensión transversal que acompaña la evaluación de las decisiones adoptadas en los diferentes niveles de automatización.

Esta incorporación permite, a su vez, conectar la responsabilidad ambiental con las tres formas de razonamiento desarrolladas previamente, de manera que cada una contribuya desde una función diferente pero complementaria. Así, el positivismo proporciona indicadores para medir el consumo de energía, las emisiones o la eficiencia de los recursos; el realismo crítico permite investigar las condiciones de diseño, producción y operación que pueden estar generando determinados impactos, mientras que la hermenéutica amplía el análisis al considerar quiénes pueden verse afectados, qué valores intervienen en la decisión y qué consecuencias pueden proyectarse hacia el futuro. Esta articulación explica su incorporación como dimensión transversal en la Tabla 4, debido a que una decisión industrial no debería responder únicamente a la pregunta de si una solución funciona, sino también a si sus resultados son ambientalmente responsables y justificables en el tiempo [11].

4.7 Interacción humano-IA y supervisión filosóficamente informada

La incorporación de esta responsabilidad amplía el criterio con el que se valoran las decisiones tecnológicas y conduce a una cuestión igualmente importante: determinar quién debe conservar la capacidad de juzgar y asumir responsabilidad cuando un sistema inteligente participa activamente en una decisión. Esta pregunta adquiere relevancia cuando la inteligencia artificial deja de limitarse al procesamiento de información y comienza a generar recomendaciones o intervenir sobre procesos industriales, porque la decisión pasa a construirse mediante la interacción entre el conocimiento profesional y las capacidades del sistema. Vallor y Vierkant [13] explican esta relación mediante una concepción de agencia distribuida entre componentes humanos y tecnológicos; sin embargo, dicha interacción no implica transferir completamente el juicio al algoritmo, pues una recomendación sustentada en grandes volúmenes de datos también puede requerir ser cuestionada cuando sus resultados no sean suficientemente explicables o no respondan adecuadamente al contexto [3, 12, 4].

Por esta razón, la supervisión humana propuesta no consiste solamente en observar lo que hace el sistema, sino en disponer de las capacidades necesarias para aceptar, cuestionar o detener una recomendación cuando las circunstancias lo justifiquen. En este proceso, el realismo crítico permite examinar las causas y estructuras que pueden explicar una respuesta algorítmica [6, 7], mientras que la hermenéutica ayuda a interpretar su significado dentro del contexto humano y organizacional [8]. De este modo, la colaboración humano-IA combina las capacidades analíticas del sistema con el conocimiento y juicio profesional [21], al tiempo que permite mantener trazabilidad sobre la responsabilidad de la decisión [22]. Por tanto, la Tabla 4, el Algoritmo 1 y la Figura 5 convergen en un marco donde cuantificación, explicación causal, interpretación contextual, responsabilidad ambiental y supervisión humana actúan de manera complementaria.

5. Discusión

A partir de los resultados obtenidos, la principal contribución del estudio radica en mostrar que la formación para la Industria 4.0 requiere ampliar el dominio técnico hacia capacidades que permitan explicar, interpretar y valorar las decisiones tomadas en sistemas inteligentes de automatización. En este sentido, el predominio del enfoque positivista resulta adecuado para desarrollar competencias de medición, programación y control, pero puede ser insuficiente cuando el profesional enfrenta problemas que no pueden comprenderse únicamente mediante indicadores cuantitativos. Esta interpretación coincide con Lysak [2], quien advierte que una formación en ingeniería sin fundamentos filosóficos suficientemente integrados puede generar limitaciones epistemológicas y éticas, especialmente ante tecnologías de mayor autonomía. Por ello, la propuesta presentada en la Tabla 4 amplía las competencias técnicas mediante capacidades causales, interpretativas y ambientales que pueden combinarse conforme aumenta la complejidad de la decisión, en correspondencia con la teoría de capacidades dinámicas [18].

Esta necesidad de ampliar el razonamiento se hace más evidente cuando una anomalía no puede explicarse mediante los datos disponibles, pues conocer qué ocurrió no significa necesariamente comprender por qué ocurrió. Frente a esta diferencia, el realismo crítico aporta la retroducción causal como una forma

de investigar los mecanismos que pueden generar un determinado evento [6, 7]. Mientras que la hermenéutica permite extender el análisis hacia la manera en que operadores y otros actores interpretan la situación dentro de su contexto. Esta complementariedad resulta consistente con los planteamientos sobre diseño centrado en el humano en Industria 4.0 y 5.0 [9, 10] y adquiere una expresión operativa en el Algoritmo 1, cuya secuencia se representa en la Figura 5. De esta manera, la propuesta conecta los fundamentos filosóficos con un procedimiento que orienta cuándo medir, cuándo buscar causas y cuándo incorporar la interpretación contextual.

Ahora bien, esta integración adquiere una importancia adicional cuando la inteligencia artificial participa directamente en la recomendación o ejecución de decisiones, ya que en estos escenarios no basta con determinar si el sistema alcanza un resultado técnicamente correcto, sino que también debe ser posible cuestionar su razonamiento y reconocer sus consecuencias. La literatura advierte que la supervisión humana puede perder efectividad cuando el profesional no dispone de conocimientos suficientes para evaluar una recomendación algorítmica y termina aceptándola de manera pasiva [3, 12, 4]. Ante este problema, el punto de supervisión humana incorporado en el Algoritmo 1 busca preservar la capacidad de juicio del profesional, mientras que la dimensión ambiental incluida en la Tabla 4 amplía ese juicio hacia las consecuencias ecológicas y futuras [11]. En efecto, ambos elementos permiten comprender la automatización como un proceso en el que eficiencia, responsabilidad y capacidad de cuestionamiento deben mantenerse relacionadas.

En este mismo camino, los resultados deben interpretarse considerando el alcance documental y conceptual del estudio, dado que las relaciones propuestas proceden del análisis y articulación de las fuentes organizadas en la Tabla 2 y no de pruebas realizadas con profesionales o sistemas industriales en funcionamiento. En consecuencia, tanto el plan integrado de capacidades de la Tabla 4 como el Algoritmo 1 constituyen propuestas conceptuales sustentadas en la literatura, pero todavía requieren ser evaluadas empíricamente para determinar su comportamiento ante situaciones reales de automatización. Esta limitación, a su vez, establece una continuidad natural para futuras investigaciones, que podrían aplicar el marco en plantas industriales, analizar su utilidad frente a diferentes tipos de anomalías y evaluar si profesionales formados bajo estas capacidades mejoran la explicación, interpretación y supervisión de decisiones automatizadas. De este modo, la validación futura permitirá determinar hasta qué punto la integración filosófica propuesta puede trasladarse de manera efectiva desde el plano conceptual hacia la práctica industrial.

6. Conclusiones

El propósito de este trabajo fue analizar la relación entre las corrientes filosóficas y la toma de decisiones en sistemas inteligentes de automatización industrial, con el fin de proponer un marco integrado de capacidades para la Industria 4.0. A partir del análisis documental, se concluye que el positivismo proporciona una base indispensable para la medición, verificación y control de los procesos, aunque su alcance resulta limitado cuando las decisiones requieren comprender causas no evidentes, interpretar situaciones humanas o responder ante sistemas inteligentes de mayor complejidad; en estos escenarios, el realismo crítico amplía el razonamiento mediante la identificación de mecanismos causales, mientras que la hermenéutica incorpora la comprensión del contexto y de las interacciones entre las personas y la tecnología. Por otro lado, la articulación de estas perspectivas permite plantear una formación profesional que no se limite al dominio técnico, sino que integre capacidades cuantitativas, causales e interpretativas para responder a problemas de diferente complejidad. A ello se suma la filosofía ambiental como dimensión transversal, mediante la cual la eficiencia tecnológica se relaciona con el uso responsable de los recursos, las consecuencias ecológicas y la responsabilidad frente a escenarios futuros. En este contexto, el estudio aporta una perspectiva transdisciplinaria para fortalecer la capacidad de los profesionales de comprender, cuestionar y supervisar decisiones en entornos industriales progresivamente automatizados, manteniendo el juicio humano como elemento necesario frente a la creciente participación de la inteligencia artificial.

Como línea de trabajo futura, se plantea avanzar desde la construcción conceptual desarrollada en este estudio hacia su validación en entornos industriales reales, mediante estudios de caso o procesos de investigación-acción con profesionales vinculados a la automatización. Esta etapa permitiría evaluar si las capacidades propuestas contribuyen efectivamente a reconocer anomalías, comprender sus causas,

interpretar situaciones de interacción humano-máquina y fortalecer la supervisión de decisiones apoyadas por sistemas inteligentes; para ello, podrían emplearse entrevistas, observación directa y otros instrumentos que permitan contrastar el marco con la experiencia de los profesionales y con situaciones concretas de planta. Asimismo, futuras investigaciones podrían ampliar el corpus documental mediante la incorporación de otras bases de datos, idiomas y contextos industriales, con el propósito de determinar si las relaciones filosóficas identificadas mantienen su pertinencia en diferentes tradiciones de formación en ingeniería y entornos tecnológicos. De esta manera, sería posible fortalecer progresivamente la validez práctica y la capacidad de generalización del marco propuesto.

Contribución de los autores. Los autores confirman su contribución al artículo de la siguiente manera: conceptualización, D.C.; metodología, D.C.; investigación, V.M.; análisis formal, V.M.; redacción – preparación del borrador original, D.C.; redacción – revisión y edición, B.L.G.P.; visualización, B.L.G.P. Todos los autores revisaron los resultados y aprobaron la versión final del manuscrito.

Agradecimientos. A la Universidad Privada Dr. Rafael Belloso Chacín por el apoyo institucional brindado para el desarrollo de esta investigación documental.

Financiamiento. El estudio ha sido autofinanciado por los autores.

Declaración de consentimiento informado. No aplica. El estudio es de naturaleza documental; no involucró la aplicación de instrumentos, entrevistas ni ningún otro procedimiento de recolección de datos a sujetos humanos.

Declaración de disponibilidad de datos. No aplica en el sentido de datos primarios, ya que el estudio no generó dichos datos. El corpus documental completo que sustenta los hallazgos está referenciado en la bibliografía y es de acceso público.

Conflictos de intereses. Los autores declaran no tener conflictos de intereses.

Declaración de uso de inteligencia artificial. Los autores emplearon herramientas de inteligencia artificial generativa exclusivamente para apoyar, la reorganización estructural del manuscrito y la mejora del estilo del texto. El análisis, la interpretación documental, el diseño del algoritmo propuesto y las conclusiones son responsabilidad intelectual exclusiva de los autores, quienes validaron y aprobaron la versión final.

Referencias

- [1] E. Y. Nakagawa, P. O. Antonino, F. Schnicke, R. Capilla, T. Kuhn, and P. Liggesmeyer, “Industry 4.0 reference architectures: State of the art and future trends,” *Computers & Industrial Engineering*, vol. 156, p. 107241, 2021. doi: 10.1016/j.cie.2021.107241.
- [2] V. Lysak, “A theoretical analysis of philosophical dimensions in engineering education,” *Journal of Contemporary Philosophical and Anthropological Studies*, vol. 2, no. 3, 2024. doi: 10.59652/jc-pas.v2i3.307.
- [3] J. Laux, “Institutionalised distrust and human oversight of artificial intelligence: Towards a democratic design of ai governance under the european union ai act,” *AI & Society*, vol. 39, no. 6, pp. 2853–2866, 2023. doi: 10.1007/s00146-023-01777-z.
- [4] M. Langer, K. Baum, and N. Schlicker, “Effective human oversight of ai-based systems: A signal detection perspective on the detection of inaccurate and unfair outputs,” *Minds and Machines*, vol. 35, no. 1, 2024. doi: 10.1007/s11023-024-09701-0.
- [5] A. Ferdman and E. Ratti, “What do we teach to engineering students: Embedded ethics, morality, and politics,” *Science and Engineering Ethics*, vol. 30, no. 1, pp. 1–26, 2024. doi: 10.1007/s11948-024-00469-1.
- [6] F. C. Mukumbang, “Retroductive theorizing: A contribution of critical realism to mixed methods research,” *Journal of Mixed Methods Research*, vol. 17, no. 1, pp. 93–114, 2023. doi: 10.1177/15586898211049847.

- [7] A. Rieder, S. Chakraborty, S. Goyal, and D. J. Berndt, "A critical realist approach to agent-based modeling: Unlocking prediction in non-positivist paradigms," *Journal of Information Technology*, vol. 40, no. 2, 2025. doi: 10.1177/02683962241280657.
- [8] A. Grunwald, A. Nordmann, and M. Sand, Eds., *Hermeneutics, History, and Technology: The Call of the Future*. London, UK: Routledge, 2023. doi: 10.4324/9781003322290.
- [9] H. Nguyen Ngoc, G. Lasa, and I. n. Iriarte, "Human-centred design in industry 4.0: Case study review and opportunities for future research," *Journal of Intelligent Manufacturing*, vol. 33, no. 1, pp. 35–76, 2022. doi: 10.1007/s10845-021-01796-x.
- [10] M. Yaqot, B. Menezes, A. Mohammed, and K. Moloney, "A state-of-the-art review and framework for human-centric automation in industry 5.0," in *Advances in Production Management Systems. Production Management Systems for Volatile, Uncertain, Complex, and Ambiguous Environments (APMS 2024), IFIP AICT vol. 729*. Springer, 2024. doi: 10.1007/978-3-031-65894-5_27.
- [11] A. R. Gammon and L. Marin, "Learning to reframe problems through moral sensitivity and critical thinking in environmental ethics for engineers," *Teaching Ethics*, vol. 22, no. 1, pp. 97–116, 2022. doi: 10.5840/tej20221013120.
- [12] S. Sterz, K. Baum, S. Biewer, H. Hermanns, A. Lauber-Rönsberg, P. Meinel, and M. Langer, "On the quest for effectiveness in human oversight: Interdisciplinary perspectives," in *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*, 2024, pp. 2495–2507. doi: 10.1145/3630106.3659048.
- [13] S. Vallor and T. Vierkant, "Find the gap: Ai, responsible agency and vulnerability," *Minds and Machines*, vol. 34, no. 3, 2024. doi: 10.1007/s11023-024-09674-0.
- [14] M. D. J. Peters, C. Marnie, H. Colquhoun, C. M. Garritty, S. Hempel, T. Horsley, E. V. Langlois, E. Lillie, K. K. O'Brien, O. Tunçalp, M. G. Wilson, W. Zarin, and A. C. Tricco, "Scoping reviews: Reinforcing and advancing the methodology and application," *Systematic Reviews*, vol. 10, p. 263, 2021. doi: 10.1186/s13643-021-01821-3.
- [15] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to conduct a bibliometric analysis: An overview and guidelines," *Journal of Business Research*, vol. 133, pp. 285–296, 2021. doi: 10.1016/j.jbusres.2021.04.070.
- [16] R. Kahlke, C. Poth, J. S. Young, and L. A. Maggio, "State of the methods for integrating qualitatively oriented critical review approaches with bibliometrics: A mixed critical bibliometric approach," *International Journal of Qualitative Methods*, vol. 25, 2026. doi: 10.1177/16094069261420676.
- [17] E. M. Martinez, P. Ponce, I. Macias, and A. Molina, "Automation pyramid as constructor for a complete digital twin, case study: A didactic manufacturing system," *Sensors*, vol. 21, no. 14, p. 4656, 2021. doi: 10.3390/s21144656.
- [18] M. Cristofaro, C. E. Helfat, and D. J. Teece, "Adapting, shaping, evolving: Refocusing on the dynamic capabilities–environment nexus," *Academy of Management Collections*, vol. 4, no. 1, pp. 20–46, 2025. doi: 10.5465/amc.2022.0008.
- [19] F. J. Folgado, D. Calderón, I. González, and A. J. Calderón, "Review of industry 4.0 from the perspective of automation and supervision systems: Definitions, architectures and recent trends," *Electronics*, vol. 13, no. 4, p. 782, 2024. doi: 10.3390/electronics13040782.
- [20] D. J. Teece, "The evolution of the dynamic capabilities framework," in *Artificiality and Sustainability in Entrepreneurship*, ser. FGF Studies in Small Business and Entrepreneurship, R. Adams, D. Grichnik, A. Pundziene, and C. Volkmann, Eds. Cham, Switzerland: Springer, 2023. doi: 10.1007/978-3-031-11371-0_6.
- [21] A. Papacharalampopoulos, P. Foteinopoulos, O. M. Karagianni, and P. Stavropoulos, "A human–ai collaborative framework for additive manufacturing modeling and decision-making," *Processes*, vol. 13, no. 12, p. 3877, 2025. doi: 10.3390/pr13123877.

- [22] H. Demirtas, “Ai responsibility gap: Not new, inevitable, unproblematic,” *Ethics and Information Technology*, vol. 27, no. 1, pp. 1–10, 2025. doi: 10.1007/s10676-024-09814-1.