

Desarrollo de una estrategia para promover

la plataforma de IA en el gobierno público

**Directrices de calidad de datos de IA en el sector público**

2024.12

**Indice**

[1. Resumen 9](#_Toc190162148)

[1.1. Objetivo 9](#_Toc190162149)

[1.2. Principios de gestión de calidad 9](#_Toc190162150)

[1.3. Ciclo de vida de los datos de entrenamiento de inteligencia artificial 10](#_Toc190162151)

[1.4. Alcance de la gestión de calidad basada en el ciclo de vida 11](#_Toc190162152)

[1.5. Procedimiento de ejecución de la inspección de calidad 12](#_Toc190162153)

[2. Marco de gestión de la calidad de los datos 14](#_Toc190162154)

[2.1. Partes interesadas en la gestión de calidad 14](#_Toc190162155)

[2.2. Estructura del marco de gestión de calidad 15](#_Toc190162156)

[3. Procedimientos y entregables de la gestión de calidad 16](#_Toc190162157)

[3.1. Fase de preparación y planificación 16](#_Toc190162158)

[3.2. Fase de implementación 17](#_Toc190162159)

[3.3. Operación y aplicación 22](#_Toc190162160)

[4. Estrategia de gestión de la calidad de los datos de inteligencia artificial 24](#_Toc190162161)

[4.1. Proceso de implementación y estrategia de gestión de la calidad de los datos 24](#_Toc190162162)

[4.2. Estrategia de gestión de la calidad en la aplicación 27](#_Toc190162163)

[5. Indicadores de gestión de calidad 29](#_Toc190162164)

[5.1. Preparación 30](#_Toc190162165)

[5.2. Integridad 37](#_Toc190162166)

[5.3. Utilidad 40](#_Toc190162167)

[5.4. Conformidad con los estándares 41](#_Toc190162168)

**Indices de figuras**

[<Figura 1 - Ciclo de vida de los datos de entrenamiento de inteligencia artificial> 10](#_Toc188635189)

[<Figura 2 – Alcance de la gestión de calidad de los datos de entrenamiento de inteligencia artificial> 11](#_Toc188635190)

[<Figura 3 – Procedimiento de inspección y mejora de la calidad de los datos de entrenamiento de inteligencia artificial> 12](#_Toc188635191)

[<Figura 4 – Partes interesadas en la gestión de calidad> 14](#_Toc188635192)

[<Figura 5 - Marco de gestión de calidad> 15](#_Toc188635193)

[<Figura 6 – Procedimiento de elaboración del plan de implementación> 16](#_Toc188635194)

[<Figura 7 – Procedimiento de planificación para la adquisición/recopilación de datos> 17](#_Toc188635195)

[<Figura 8 – Procedimiento de depuración de datos> 19](#_Toc188635196)

[<Figura 9 – Procedimiento de procesamiento de datos> 20](#_Toc188635197)

[<Figura 10 – Procedimiento de entrenamiento de datos 21> 21](#_Toc188635198)

[<Figura 11 – Procedimiento de operación y aplicación de datos 22> 22](#_Toc188635199)

[<Figura 12 – Estrategia de gestión de la calidad de los datos de inteligencia artificial 24> 24](#_Toc188635200)

[<Figura 13 – Proceso de implementación y procedimiento de gestión de la calidad de los datos 24> 24](#_Toc188635201)

[<Figura 14 – Procedimiento detallado de inspección de calidad en la fase de operación y aplicación 27> 27](#_Toc188635202)

**Indices de tablas**

[<Tabla 1 – Principios de gestión de calidad> 9](#_Toc188635208)

[<Tabla 2 – Contenidos detallados según el alcance de la gestión de calidad> 12](#_Toc188635209)

[<Tabla 3 – Definición de roles de las partes interesadas en la gestión de calidad> 14](#_Toc188635210)

[<Tabla 4 – Actividades y entregables de gestión de calidad en la fase de planificación e implementación> 17](#_Toc188635211)

[<Tabla 5 – Actividades y entregables de gestión de calidad en la fase de adquisición/recopilación de datos> 18](#_Toc188635212)

[<Tabla 6 – Actividades y entregables de gestión de calidad en la fase de depuración de datos> 20](#_Toc188635213)

[<Tabla 7 – Actividades y entregables de gestión de calidad en la fase de procesamiento de datos> 21](#_Toc188635214)

[<Tabla 8 – Actividades y entregables de gestión de calidad en la fase de aprendizaje de datos> 22](#_Toc188635215)

[<Tabla 9 – Actividades y entregables de gestión de calidad en la fase de operación y aplicación de datos> 23](#_Toc188635216)

[<Tabla 10 – Selección de datos sujetos a inspección: contenido y ejemplos> 25](#_Toc188635217)

[<Tabla 11 – Contenido y ejemplos de inspección en la fase de adquisición/recopilación de datos> 25](#_Toc188635218)

[<Tabla 12 – Contenido y ejemplos de inspección en la fase de adquisición/recopilación de datos> 26](#_Toc188635219)

[<Tabla 13 – Contenido y ejemplos de inspección en la fase de modelado> 26](#_Toc188635220)

[<Tabla 14 – Contenido y ejemplos de inspección exhaustiva> 26](#_Toc188635221)

[<Tabla 15 – Procedimiento de definición de datos sujetos a inspección> 27](#_Toc188635222)

[<Tabla 16 – Procedimiento de ejecución de la inspección de calidad> 28](#_Toc188635223)

[<Tabla 17 – Procedimiento de análisis de resultados de inspección> 28](#_Toc188635224)

[<Tabla 18 – Procedimiento de implementación de mejoras> 28](#_Toc188635225)

[<Tabla 19 – Procedimiento de control de calidad> 29](#_Toc188635226)

[<Tabla 20 – Métodos y criterios de inspección por indicador de calidad> 30](#_Toc188635227)

[<Tabla 21 – Criterios y métodos de inspección de preparación procedimental en la planificación> 31](#_Toc188635228)

[<Tabla 22 – Contenido de inspección de preparación procedimental en la planificación> 32](#_Toc188635229)

[<Tabla 23 – Criterios y métodos de inspección de preparación organizativa en la planificación> 33](#_Toc188635230)

[<Tabla 24 – Contenido de inspección de preparación organizativa en la planificación> 34](#_Toc188635231)

[<Tabla 25 – Criterios y métodos de inspección de preparación de herramientas en la planificación 34](#_Toc188635232)

[<Tabla 26 – Contenido de inspección de preparación de herramientas en la planificación> 35](#_Toc188635233)

[<Tabla 27 – Criterios y métodos de inspección de gestión de riesgos en la planificación> 35](#_Toc188635234)

[<Tabla 28 – Contenido de inspección de gestión de riesgos en la planificación> 36](#_Toc188635235)

[<Tabla 29 – Criterios y métodos de inspección de cumplimiento de seguridad en la conformidad del sistema> 36](#_Toc188635236)

[<Tabla 30 – Contenido de inspección de cumplimiento de seguridad en la conformidad del sistema> 37](#_Toc188635237)

[<Tabla 31 – Criterios y métodos de inspección de cumplimiento normativo y legal en la conformidad del sistema> 37](#_Toc188635238)

[<Tabla 32 – Contenido de inspección de cumplimiento normativo y legal en la conformidad del sistema> 38](#_Toc188635239)

[<Tabla 33 – Criterios y métodos de inspección de integridad en la recopilación de datos> 39](#_Toc188635240)

[<Tabla 34 – Contenido de inspección de integridad en la recopilación de datos> 39](#_Toc188635241)

[<Tabla 35 – Criterios y métodos de inspección de integridad en la depuración de datos> 39](#_Toc188635242)

[<Tabla 36 – Contenido de inspección de integridad en la depuración de datos> 40](#_Toc188635243)

[<Tabla 37 – Criterios y métodos de inspección de integridad en el procesamiento de datos> 40](#_Toc188635244)

[<Tabla 38 – Contenido de inspección de integridad en el procesamiento de datos> 41](#_Toc188635245)

[<Tabla 39 – Criterios y métodos de inspección de usabilidad> 41](#_Toc188635246)

[<Tabla 40 – Contenido de inspección de usabilidad> 41](#_Toc188635247)

[<Tabla 41 – Criterios y métodos de inspección de flexibilidad> 42](#_Toc188635248)

[<Tabla 42 – Contenido de inspección de flexibilidad> 42](#_Toc188635249)

[<Tabla 43 – Criterios y métodos de inspección de diversidad> 42](#_Toc188635250)

[<Tabla 44 – Contenido de inspección de diversidad> 43](#_Toc188635251)

[<Tabla 45 – Criterios y métodos de inspección de fiabilidad> 43](#_Toc188635252)

[<Tabla 46 – Contenido de inspección de fiabilidad> 43](#_Toc188635253)

[<Tabla 47 – Criterios y métodos de inspección de suficiencia> 44](#_Toc188635254)

[<Tabla 48 – Contenido de inspección de suficiencia> 44](#_Toc188635255)

[<Tabla 49 – Criterios y métodos de inspección de uniformidad> 44](#_Toc188635256)

[<Tabla 50 – Contenido de inspección de uniformidad> 45](#_Toc188635257)

[<Tabla 51 – Criterios y métodos de inspección de veracidad> 45](#_Toc188635258)

[<Tabla 52 – Contenido de inspección de veracidad> 46](#_Toc188635259)

[<Tabla 53 – Criterios y métodos de inspección de equidad> 46](#_Toc188635260)

[<Tabla 54 – Contenido de inspección de equidad> 46](#_Toc188635261)

**Lista de abreviaturas**

| Acrónimos | Definición |
| --- | --- |
| AI | Artificial Intelligence |
| AI RMF | Artificial Intelligence Risk Management Framework |
| API | Application Programming Interface |
| CNN | Convolutional Neural Network |
| CSP | Cloud Service Provider |
| D-Gov | Digital Government |
| DB | Database |
| DL | Deep Learning |
| DLC | Data Life Cycle |
| DLCF | Data Life Cycle Framework |
| DQMLC | Data Quality Management Life Cycle |
| DQPF | Data Quality Process Framework |
| GAN | Generative Adversarial Network |
| HPC | High-Performance Computing |
| ISO/IEC | International Organization for Standardization / International Electrotechnical Commission |
| ITU | International Telecommunication Union |
| KPI | Key Performance Indicator |
| ML | Machine Learning |
| MSP | Managed Service Provider |
| NIA DGCC | National Information Society Agency Digital Government Cooperation Center |
| NIST | National Institute of Standards and Technology |
| NLP | Natural Language Processing |
| ONP | Office of National Pension (Peru) |
| PCM | Presidency of the Council of Ministers (Peru) |
| PIDE | Digital Government Platform (Peru) |
| PNGD | Plataforma Nacional de Gobierno Digital |
| PNTD | Plan Nacional de Transformación Digital |
| RDBMS | Relational Database Management System |
| RFP | Request for Proposal |
| SEGDI | Secretaría de Gobierno Digital |
| SGTD | Secretaría de Gobierno y Transformación Digital |
| SLA | Service Level Agreement |
| UI | User Interface |

* + - * 1. Resumen
  1. Objectivo

La guía de calidad de datos proporciona orientación para facilitar la comprensión y el uso adecuado y responsable de la gestión de calidad de los datos de inteligencia artificial en el gobierno y en el sector público en general. Las instituciones deben trabajar en conjunto con el marco a lo largo de todo el proceso de planificación y evaluación de proyectos de inteligencia artificial, garantizando la revisión periódica de la calidad de los datos. Esta guía establece estándares que refuerzan la gestión de calidad de los datos de inteligencia artificial mediante indicadores universales y comunes, teniendo en cuenta las características de recopilación, procesamiento y aplicación de diversos conjuntos de datos de entrenamiento en inteligencia artificial. Además, se establecen criterios para permitir una gestión de calidad continua y sostenible.

* 1. Principios de gestión de calidad

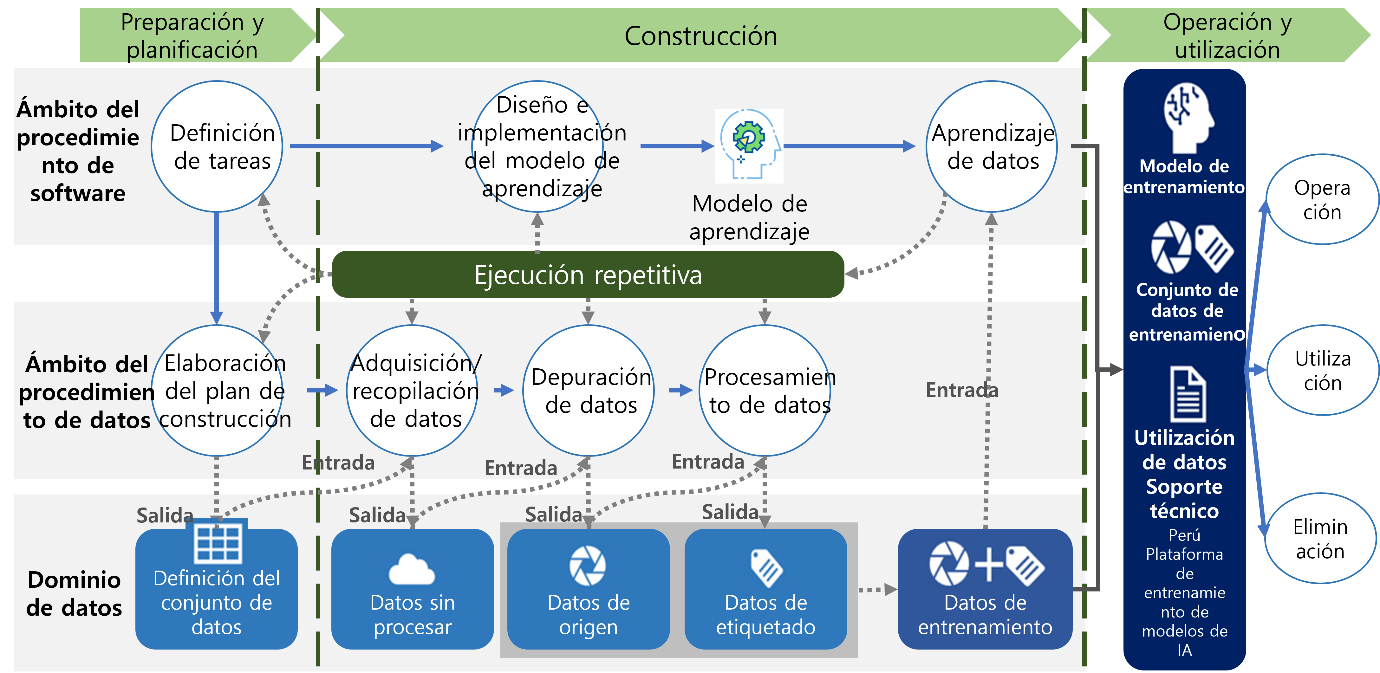
Para garantizar la adquisición de datos de entrenamiento de inteligencia artificial de alta calidad, es fundamental asegurar la calidad de los distintos tipos de datos de entrenamiento, incluidos los datos en bruto, los datos fuente y los datos etiquetados. Asimismo, se requiere el establecimiento de un sistema de gestión de calidad que permita mantener o mejorar continuamente la calidad de los datos en función de las necesidades de los usuarios. Además, es necesario definir principios de gestión de calidad de alto nivel que orienten la gestión de calidad de los datos de entrenamiento de inteligencia artificial.

|  |  |
| --- | --- |
| Clasificación | Principios |
| Aspecto de gestión de calidad | Se debe garantizar la calidad a lo largo de todo el ciclo de vida de los datos de entrenamiento de inteligencia artificial. |
| La mejora de la calidad de los datos de entrenamiento de inteligencia artificial debe ser continua y permanente. |
| Se debe establecer una organización para la gestión de calidad de los datos, asegurando la ejecución de actividades conforme a los roles y responsabilidades definidos. |
| Se debe contar con un sistema de apoyo, incluyendo formación en gestión de calidad, para fortalecer las capacidades de la organización en este ámbito. |
| Aspecto de los datos | Se deben cumplir los requisitos necesarios para el entrenamiento de los datos de inteligencia artificial. |
| Los datos deben estar disponibles para su uso sin restricciones legales. |
| Los datos deben ser adquiridos y recopilados de manera acorde con el propósito del aprendizaje, asegurando su depuración sin duplicaciones y conforme al objetivo deseado. |
| Se debe garantizar la precisión de la información de etiquetado, como los valores de referencia (Ground Truth), asignados a los datos de entrenamiento de inteligencia artificial. |
| La efectividad del modelo de inteligencia artificial debe ser validada y alineada con el objetivo establecido. |
| Para lograr una inteligencia artificial confiable, los datos de entrenamiento no deben presentar problemas éticos. |

<Tabla 1 – Principios de gestión de calidad>

* 1. Ciclo de vida de los datos de entrenamiento de inteligencia artificial

Para identificar las actividades de gestión de calidad de los datos de inteligencia artificial, es necesario analizar su ciclo de vida y establecer categorías de gestión de calidad basadas en este. El ciclo de vida de los datos de inteligencia artificial, desde su creación hasta su apertura para uso público, se desarrolla de la siguiente manera.

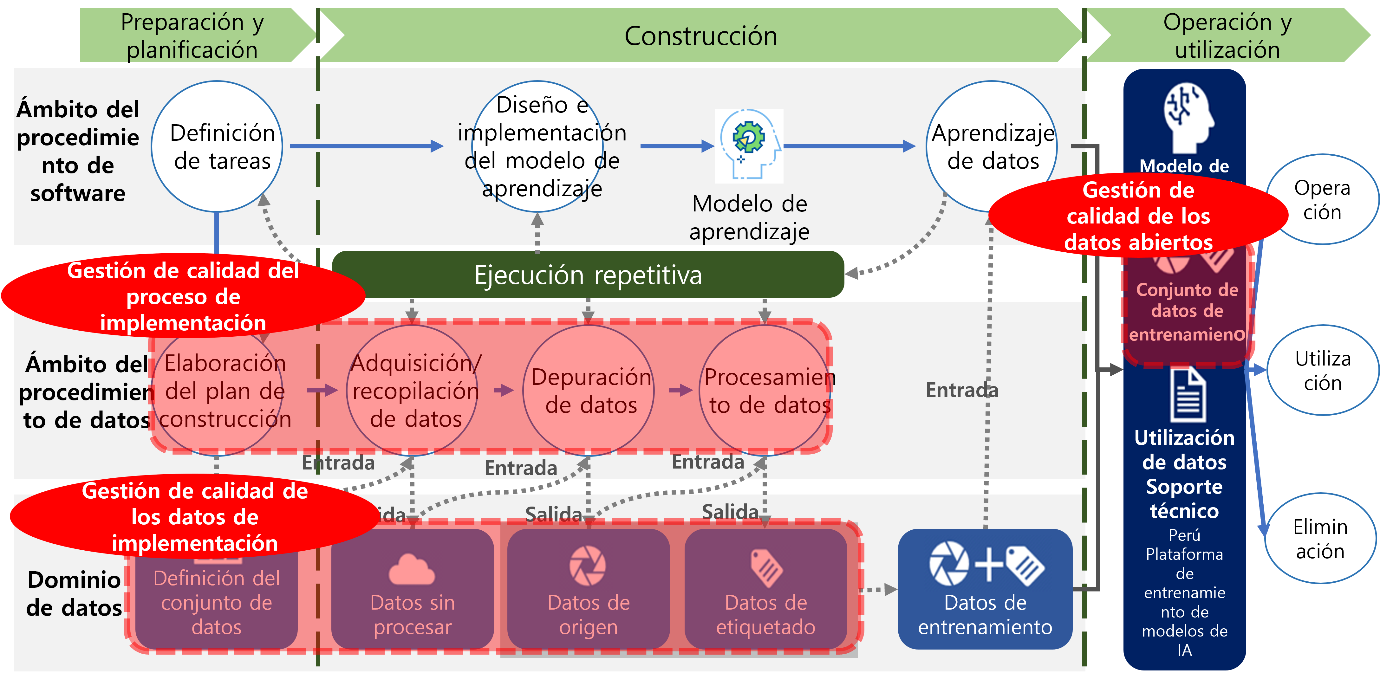


<Figura 1 - Ciclo de vida de los datos de entrenamiento de inteligencia artificial>

El proceso de construcción de datos se divide en tres grandes áreas: "Preparación y planificación", "Construcción" y "Operación y aplicación". En cada una de estas áreas, se identifican los responsables de la gestión de calidad y se definen actividades específicas para evitar dificultades de comunicación o confusión durante la ejecución de las tareas.

* **Fase de preparación y planificación**
* Consiste en definir con claridad el problema que la inteligencia artificial busca resolver mediante el aprendizaje automático y en diseñar y especificar los datos de entrenamiento de inteligencia artificial necesarios para dicha solución. Durante la planificación de la construcción, se establecen aspectos clave como el "Resumen de construcción", "Construcción de datos", "Plan de inspección de calidad" y "Definición del conjunto de datos".
* **Fase de construcción**
* Se lleva a cabo la generación directa de los datos necesarios para el aprendizaje automático de la inteligencia artificial o la adquisición de estos desde organizaciones o sistemas existentes, asegurando la obtención de los "datos en bruto".
* Los datos en bruto adquiridos se ajustan al formato y tamaño requeridos para el aprendizaje automático, eliminando posibles duplicaciones.
* Durante la obtención de los datos en bruto, se realiza un proceso de anonimización para garantizar la protección de datos personales, asegurando así la obtención de los "datos fuente".
* Se lleva a cabo el proceso de "etiquetado", en el cual se asignan etiquetas adecuadas a los datos fuente de acuerdo con su función y propósito, permitiendo su uso en el aprendizaje automático de inteligencia artificial. Los "datos etiquetados" generados en este proceso incluyen el "valor de referencia" (Ground Truth) asignado a los datos fuente, así como información sobre el formato de archivo, resolución y otras características de los datos, que se organizan en un conjunto de "anotaciones".
* **Fase de operación y aplicación**
* Se utiliza el "conjunto de datos de entrenamiento" para entrenar el modelo de inteligencia artificial previamente definido. Además, se llevan a cabo actividades para mejorar o ajustar el rendimiento del modelo de inteligencia artificial entrenado.
  1. Alcance de la gestión de calidad basada en el ciclo de vida

El alcance de la gestión de calidad, basado en el ciclo de vida de los datos de entrenamiento de inteligencia artificial, se define de la siguiente manera.



<Figura 2 – Alcance de la gestión de calidad de los datos de entrenamiento de inteligencia artificial>

* **Gestión de calidad basada en el proceso de construcción**
* Monitoreo de las actividades para garantizar el nivel de calidad deseado en el proceso de construcción de datos, incluyendo la adquisición/recopilación, depuración y procesamiento de datos. Además, se realizan mejoras en función de los problemas identificados desde la perspectiva de la gestión de calidad.
* **Gestión de calidad de los datos construidos**
* Inspección de la calidad de los datos generados a través del proyecto de construcción, incluyendo los datos en bruto, datos fuente, datos etiquetados y datos no etiquetados. Asimismo, se llevan a cabo acciones correctivas para solucionar errores detectados.
* **Gestión de calidad de los datos abiertos**
* Ejecución de actividades continuas para mejorar la calidad del conjunto de datos de entrenamiento en la fase de operación y aplicación, centrándose en los datos cargados en la plataforma de aprendizaje del modelo de IA en Perú. Asimismo, en caso de recibir sugerencias de mejora sobre la calidad de los datos de entrenamiento abiertos al sector privado, se responde y se toman medidas proactivamente.

|  |  |  |
| --- | --- | --- |
| Fase | Alcance de la gestión de calidad | Contenido de ejecución |
| Fase de preparación y planificación | Proceso de construcción | * La entidad ejecutora del proyecto lleva a cabo la gestión de calidad de manera autónoma. * En caso de que se presenten sugerencias de mejora sobre la calidad del conjunto de datos de entrenamiento abiertos al sector privado, la entidad ejecutora del proyecto realiza directamente las actividades de mejora de calidad. |
| Fase de construcción | Datos construidos |
| Fase de operación y aplicación | Datos abiertos | * Para garantizar la gestión continua de la calidad de los datos abiertos a través de la plataforma de IA compartida de Perú, la entidad ejecutora del proyecto y la entidad gestora del proyecto llevan a cabo conjuntamente actividades de gestión de calidad. * La entidad gestora del proyecto revisa las sugerencias de mejora de calidad de los datos recibidas de los usuarios y solicita a la entidad ejecutora del proyecto la implementación de mejoras en los datos correspondientes. La entidad ejecutora, a su vez, refleja estas solicitudes y lleva a cabo las acciones necesarias para mejorar la calidad. |

<Tabla 2 – Contenidos detallados de ejecución según el alcance de la gestión de calidad>

* 1. Procedimiento de ejecución de la inspección de calidad

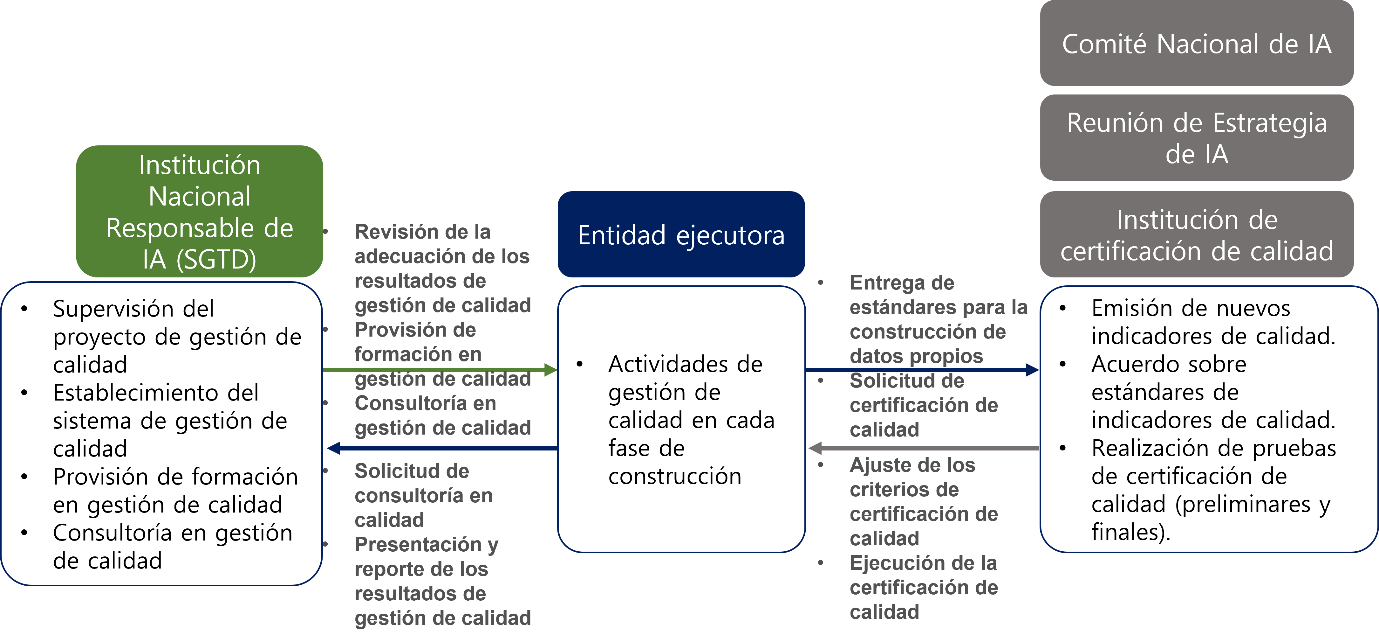
Los responsables de la gestión de calidad de la entidad ejecutora del proyecto y las instituciones participantes deben seguir el siguiente procedimiento general de inspección y mejora de calidad para la gestión de calidad en cada fase del proceso de construcción.



<Figura 3 – Procedimiento de inspección y mejora de la calidad de los datos de entrenamiento de inteligencia artificial>

* **Definición de los datos sujetos a inspección**
* Selección de los datos en bruto, datos fuente y datos etiquetados como objetos de inspección, y establecimiento de un plan de inspección de calidad para determinar si dichos datos cumplen con el nivel de calidad requerido.
* **Ejecución de la inspección de calidad**
* Ejecución de la inspección de calidad aplicando técnicas de verificación, como listas de control, a los datos definidos como objetos de inspección, evaluando su conformidad con los indicadores de gestión de calidad, incluyendo adecuación, exactitud, validez, preparación, integridad y utilidad.
* **Análisis de los resultados de la inspección**
* Etapa en la que, con base en los resultados de la inspección de calidad, se identifican los principales problemas de calidad, se analizan sus causas fundamentales y se generan oportunidades de mejora para su resolución.
* **Implementación de mejoras**
* Definición de un plan y estrategias de mejora para la resolución de problemas de calidad, determinación de prioridades y ejecución de actividades de mejora de calidad en cada área, como la corrección de datos y trabajos adicionales, según el orden de prioridad establecido.
* **Control de calidad**
* Etapa en la que se verifican y revisan los resultados de la implementación de mejoras, permitiendo redefinir los objetivos de calidad y llevar a cabo actividades de gestión de calidad para prevenir la recurrencia de problemas y mantener datos de alta calidad.
  + - * 1. Marco de gestión de la calidad de los datos
  1. Partes interesadas en la gestión de calidad

Las partes interesadas en la gestión de calidad de los datos de entrenamiento de inteligencia artificial están conformadas por: la entidad ejecutora, que lleva a cabo las actividades de gestión de calidad en cada fase de construcción; la agencia nacional de IA SGTD, que supervisa el programa nacional de gestión de calidad de datos; y la entidad de verificación de calidad, bajo el Consejo de Estrategia de IA, encargada de la verificación de calidad por terceros.

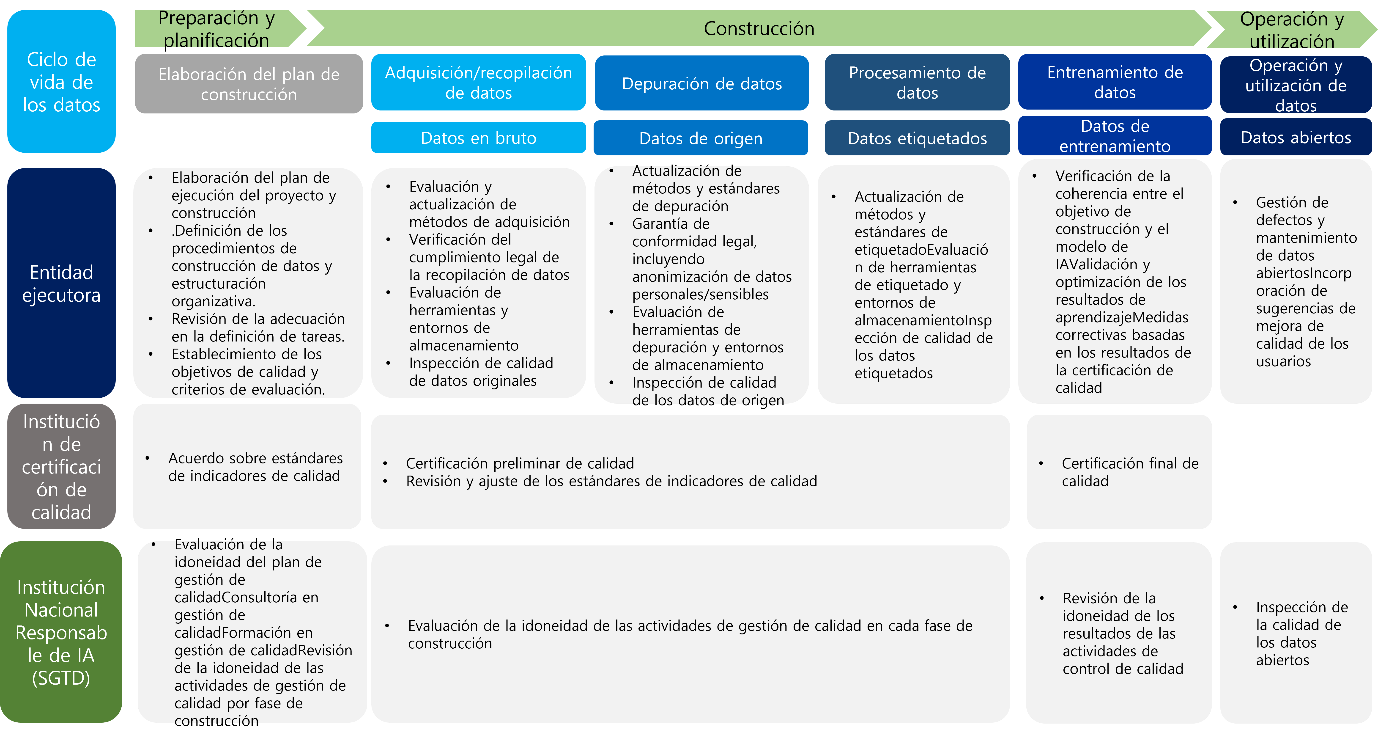


<Figura 4 – Partes interesadas en la gestión de calidad>

|  |  |
| --- | --- |
| Parte interesada | Rol |
| Entidad ejecutora | * Durante la fase de planificación, se establece una organización de calidad para llevar a cabo actividades de gestión de calidad en cada etapa. |
| Agencia nacional de IA (SGTD) | * La Dirección de Transformación Digital (SGTD) de la Oficina del Primer Ministro supervisa el programa de gestión de calidad, brinda apoyo a las entidades para que cumplan adecuadamente sus funciones en gestión de calidad, y se encarga de la gestión de los entregables finales, la consultoría y el fortalecimiento de capacidades en gestión de calidad. |
| Entidad de verificación de calidad | * La entidad de verificación de calidad, bajo el Consejo de Estrategia de IA de la Comisión Nacional de IA, acuerda con la entidad ejecutora los indicadores de calidad y realiza la verificación de calidad con base en los indicadores acordados, asegurando la validación de los resultados de la verificación. |

<Tabla 3 – Definición de roles de las partes interesadas en la gestión de calidad>

* 1. Estructura del marco de gestión de calidad

El marco de gestión de calidad necesario para la ejecución del proceso en las fases de construcción y operación de los datos de inteligencia artificial se define en función del ciclo de vida de los datos, abarcando las fases de preparación y planificación, construcción, y operación y aplicación. En cada una de estas fases, se establecen las actividades de gestión de calidad a ser ejecutadas por la entidad ejecutora, la entidad de verificación de calidad y la agencia nacional de IA (SGTD), según se detalla a continuación.

<Figura 5 - Marco de gestión de calidad>

* + - * 1. Procedimientos y entregables de la gestión de calidad

En los procedimientos de gestión de calidad, se definen las actividades específicas de gestión de calidad, los entregables requeridos y los aspectos a considerar en cada fase del ciclo de vida de los datos, incluyendo preparación y planificación, construcción, y operación y aplicación.

* 1. Fase de preparación y planificación
     1. Elaboración del plan de construcción

En esta etapa, se presentan los procedimientos y componentes necesarios para la construcción de datos según cada tipo de datos de entrenamiento de inteligencia artificial. Se diseña el conjunto de datos de entrenamiento de acuerdo con el propósito del proyecto, incluyendo la preparación de materiales y la redacción del plan de ejecución del proyecto. A través de la elaboración de un documento de planificación de construcción, que puede actualizarse y ajustarse hasta la finalización del proyecto, se busca minimizar errores durante el proceso de construcción de datos, garantizando la coherencia y alineación con los objetivos del proyecto.



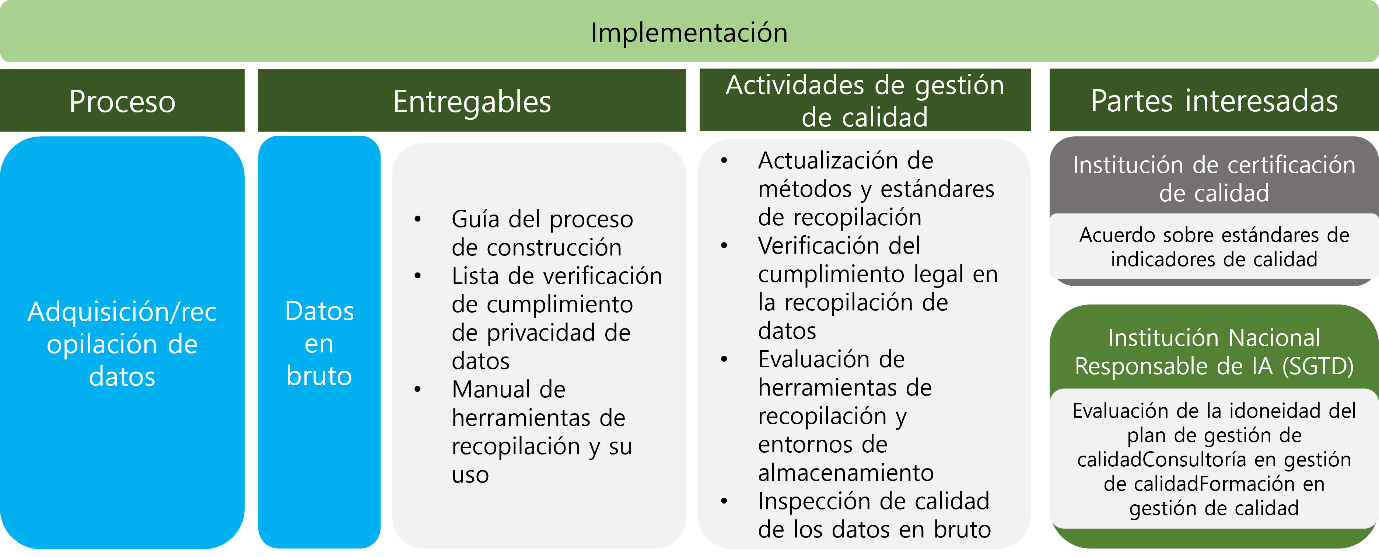
<Figura 6 – Procedimiento de elaboración del plan de construcción>

|  |  |  |  |
| --- | --- | --- | --- |
| Actividades de gestión de calidad | Contenido principal | Entregables | Definición de entregables |
| Elaboración del plan de ejecución del proyecto y construcción | * Revisión del establecimiento del propósito y alcance de la construcción, así como del plan de ejecución. * Revisión del plan de gestión de calidad. * Evaluación de la adecuación del nivel de objetivos de calidad y revisión de la definición de los estándares de gestión de calidad. | Plan de ejecución del proyecto  Plan de construcción  Plan de gestión de calidad | * Plan de ejecución del proyecto: Documento oficial que detalla el plan de ejecución presentado durante la convocatoria del proyecto. * Plan de construcción: Incluye el proceso de construcción de datos, la definición de los datos construidos y los procedimientos de inspección de datos. * Plan de gestión de calidad: Contempla los indicadores y estándares de calidad, los procedimientos de inspección y la formación en calidad. * Revisión y modificación del plan de construcción y del plan de gestión de calidad según las opiniones del Comité de Coordinación del Proyecto y los acuerdos sobre los indicadores de calidad. |
| Procedimientos de construcción de datos, estructura organizativa | * Definición de la especificación de los datos y su estructura. * Revisión del establecimiento de roles y responsabilidades en los procedimientos, organización y herramientas para la construcción de datos. * Revisión del resumen del conjunto de datos, detalles de su composición, criterios de diseño, distribución, estructura, proceso de construcción, inspección y actividades de gestión de calidad. |
| Revisión de la adecuación en la definición de funciones | * Revisión del plan para garantizar la coherencia y alineación con el propósito de construcción. | Guía del proceso de construcción | * Redacción de un borrador de definición de tareas y guía por proceso en la fase de planificación de la construcción para garantizar la coherencia y alineación con el propósito de construcción. * Actualización y gestión de cambios en los entregables según cada etapa del proceso de construcción. |
| Establecimiento de objetivos de calidad y criterios de verificación | * Establecimiento de los criterios de indicadores de calidad basados en los documentos de verificación, como el plan de ejecución del proyecto presentado por la entidad ejecutora, y acuerdo sobre los indicadores y objetivos de verificación de calidad. * (Entidad ejecutora ↔ Entidad de verificación de calidad, Agencia nacional de IA) | Documento de estándares de indicadores de calidad | * Entregables esenciales para la verificación de calidad, acordados con la entidad de verificación de calidad después de la formalización del convenio. * En la presentación final, se adjunta como anexo al plan de gestión de calidad dentro del plan de ejecución del proyecto. |

<Tabla 4 – Actividades de gestión de calidad y entregables en la fase de elaboración del plan de construcción>

* 1. Fase de construcción
     1. Adquisición/Recopilación de datos

La fase de construcción de datos, en la que se recopilan datos del mundo real o se reutilizan datos existentes para obtener datos en bruto adecuados al propósito del proyecto, es una etapa crucial que determina el nivel de calidad general de los datos de entrenamiento de inteligencia artificial. Esto se debe a la necesidad de cumplir con los requisitos del aprendizaje de IA y al hecho de que los datos construidos serán posteriormente abiertos al sector privado. Por lo tanto, es fundamental identificar de antemano los requisitos de las etapas de depuración y procesamiento, así como establecer un entorno de datos adecuado para los objetivos de etiquetado. Durante la fase de construcción, no solo se debe gestionar la calidad de los datos en sí, sino también la calidad del proceso de generación de datos.



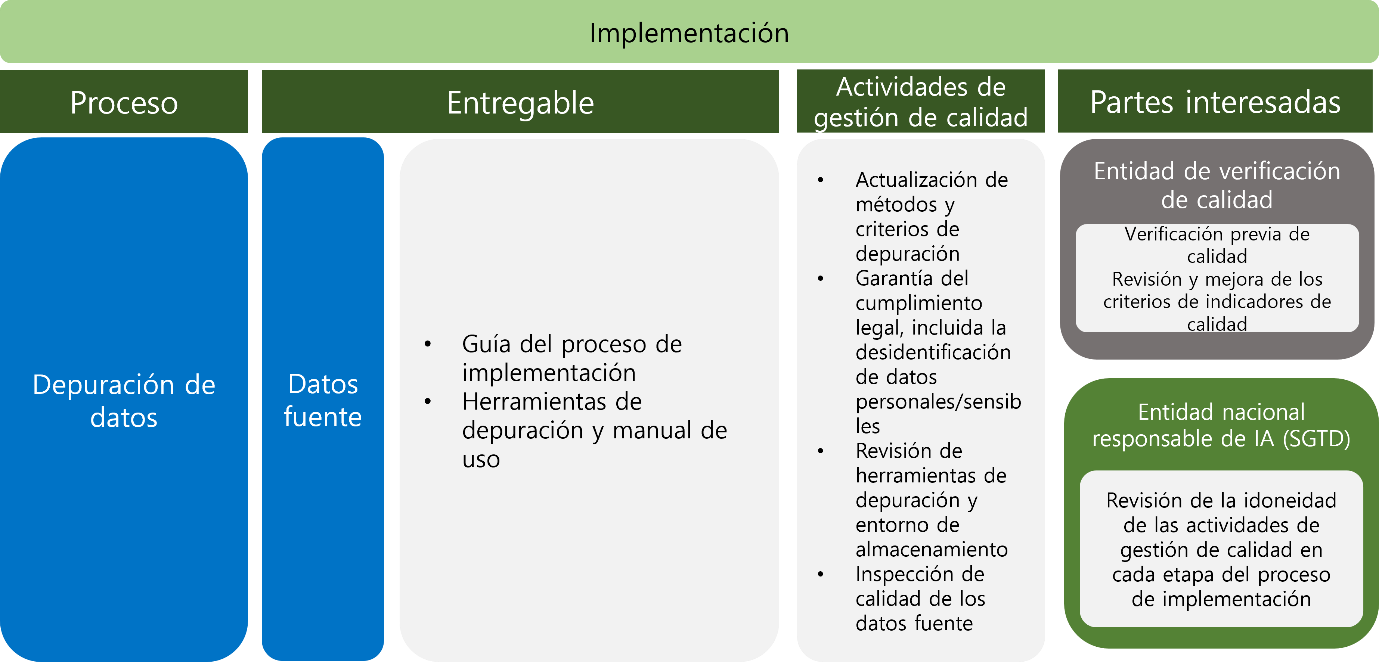
<Figura 7 – Procedimiento de planificación para la adquisición/recopilación de datos>

|  |  |  |  |
| --- | --- | --- | --- |
| Actividades de gestión de calidad | Contenido principal | Entregables | Definición de entregables |
| Métodos y criterios actualizados de adquisición/recopilación | * Métodos, procedimientos y adecuación legal y normativa para la adquisición y obtención de datos. * Establecimiento de estándares de calidad, organización, procedimientos y herramientas para la fase de adquisición/recopilación en relación con los elementos de datos y la cantidad a construir. * Desarrollo de un sistema para los métodos y criterios de adquisición de datos, formación y verificación. * Gestión de cambios en los criterios de adquisición de datos. | Guía del proceso de construcción (Adquisición/Recopilación) | * Presentación de una guía (directriz) que define los objetivos, procedimientos, criterios, organización, métodos y herramientas de adquisición/recopilación de datos, así como los procesos de inspección. |
| Revisión de la base legal para la recopilación de datos | * Obtención de documentación justificativa del cumplimiento legal y normativo (protección de datos personales, derechos de imagen, derechos de autor). * Establecimiento de medidas para prevenir sesgos en la adquisición/recopilación de datos. | Lista de verificación de cumplimiento de protección de datos personales | * Lista de verificación para evaluar si la adquisición/recopilación de datos se lleva a cabo conforme a la Ley de Protección de Datos Personales, derechos de autor, propiedad intelectual y derechos de imagen. |
| Evaluación de herramientas de recopilación y entorno de almacenamiento | * Adquisición de herramientas para la recopilación/adquisición de datos y definición de su método de uso. * Establecimiento del entorno de almacenamiento para los datos recopilados/adquiridos | Herramientas de recopilación de datos y manual de uso | * Equipos o herramientas necesarias para la adquisición/recopilación de datos, como cámaras, videos, micrófonos, herramientas de rastreo (crawling), sensores, aplicaciones de recopilación, entre otros. * Manual de uso de las herramientas de recopilación y especificaciones del entorno de almacenamiento. |
| Inspección de calidad de los datos en bruto | * Verificación de que los datos en bruto cumplen con los estándares requeridos para resolver problemas mediante inteligencia artificial, incluyendo diversidad, fiabilidad, suficiencia, uniformidad, veracidad y equidad. * Comprobación de la compatibilidad técnica en términos de formato de archivo, resolución, velocidad de fotogramas, entre otros. * Análisis de la distribución de clases, la distribución de instancias y la distribución estadística de los elementos analizables en los datos. | Datos en bruto | * Datos generados en la fase de adquisición/recopilación. |
| Informe de resultados de gestión de calidad | * Elaboración de los estándares de inspección de calidad, estructura organizativa, procedimientos y resultados de la inspección. |
| Herramientas de inspección de calidad | * Definición de herramientas para la verificación de la idoneidad de los datos adquiridos, con énfasis en su conformidad con los estándares, compatibilidad técnica y diversidad estadística. * Las herramientas de inspección de calidad pueden incluir aplicaciones o listas de verificación (como formularios de inspección visual, entre otros). |

<Tabla 5 - Actividades de gestión de calidad y entregables en la fase de adquisición/recopilación de datos>

* + 1. Procedimiento de depuración de datos

Proceso en el que los datos en bruto adquiridos se depuran para adaptarlos al aprendizaje automático y se anonimizan para garantizar la protección de datos personales, obteniendo así los datos fuente. Durante esta fase, se establecen criterios de depuración claros y alineados con el propósito de construcción, permitiendo la eliminación de datos innecesarios y la ejecución de un proceso de preprocesamiento de datos de manera estructurada.



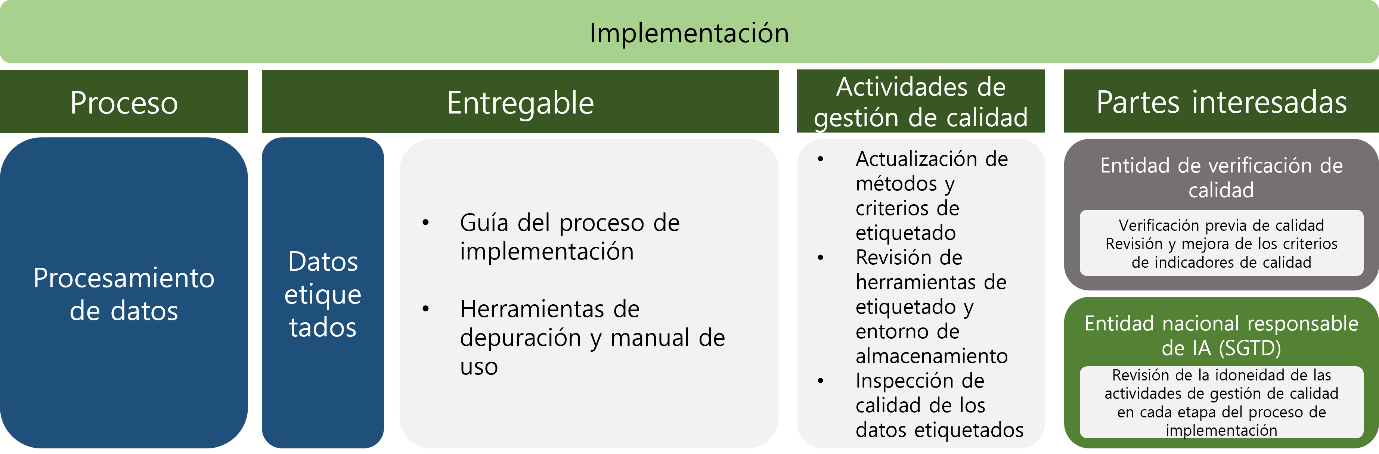
<Figura 8 – Procedimiento de depuración de datos>

|  |  |  |  |
| --- | --- | --- | --- |
| Actividades de gestión de calidad | Contenido principal | Entregables | Definición de entregables |
| Actualización de métodos y criterios de depuración | * Establecimiento de criterios de depuración para la selección adecuada de datos según el propósito de construcción. * Definición de estándares de calidad, estructura organizativa, procedimientos y herramientas para la fase de depuración. * Capacitación y formación sobre los métodos de depuración de datos y los resultados obtenidos. * Establecimiento de procedimientos y criterios de inspección para la verificación de los resultados de la depuración de datos. * Gestión de cambios en los criterios de depuración de datos. | Guía del proceso de construcción (Depuración) | * Definición del formato de los datos fuente, el alcance de la depuración, los procedimientos, criterios, organización, métodos y herramientas utilizados en la fase de depuración de datos, y presentación de una guía detallada. |
| Garantía del cumplimiento legal en la anonimización de datos personales/sensibles | * Garantía del cumplimiento legal mediante la definición de criterios y estrategias de anonimización para la protección de datos personales. |
| Revisión de herramientas de depuración y entorno de almacenamiento | * Adquisición de herramientas de edición para la fase de depuración y definición de su método de uso. * Establecimiento del entorno de almacenamiento para los datos depurados. | Herramientas de edición y manual de uso | * Herramientas para la depuración de datos, incluyendo eliminación de duplicados y anonimización de datos personales en la adquisición de datos en bruto. * Manual de uso de las herramientas de depuración y especificaciones del entorno de almacenamiento. |
| Inspección de calidad de los datos fuente | * Inspección de calidad para verificar si los datos son adecuados para su uso. * Herramientas de inspección y resultados de la verificación de calidad de los datos fuente. * Medidas correctivas basadas en los resultados de la inspección de calidad. | Datos fuente | * Datos depurados a través de procesos de preprocesamiento necesarios antes de su utilización en la fase de etiquetado, sin que aún se les haya asignado etiquetas. |
| Informe de resultados de gestión de calidad | * Elaboración de los estándares de inspección de calidad, estructura organizativa, procedimientos y resultados de la inspección. |
| Herramientas de inspección de calidad | * Definición de herramientas para la verificación de la idoneidad y duplicación de los datos adquiridos. * Las herramientas de inspección de calidad pueden incluir aplicaciones o listas de verificación (como formularios de inspección visual, entre otros). |

<Tabla 6 – Actividades de gestión de calidad y entregables en la fase de depuración de datos>

* + 1. Procesamiento de datos

Proceso en el que se asignan etiquetas a los datos fuente de acuerdo con la función y el propósito del modelo de IA, incluyendo información como el nombre del elemento, tipo y obligatoriedad. Se construye la información de anotación siguiendo una terminología estandarizada y un sistema de clasificación unificado. Además, se utilizan diversas herramientas de etiquetado para generar entregables en formatos estandarizados (CSV, JSON, etc.) y se lleva a cabo una inspección de calidad.



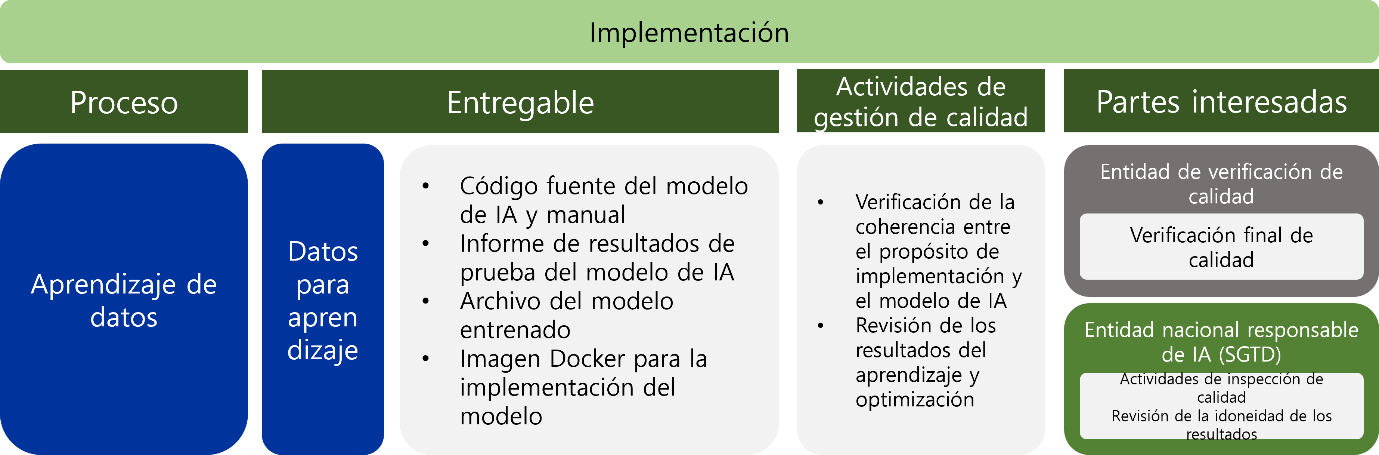
<Figura 9 – Procedimiento de procesamiento de datos>

|  |  |  |  |
| --- | --- | --- | --- |
| Actividades de gestión de calidad | Contenido principal | Entregables | Definición de entregables |
| Actualización de métodos y criterios de etiquetado | * Definición de métodos y criterios de etiquetado. * Configuración de las tareas de etiquetado y anotación necesarias para el entrenamiento. * Establecimiento de criterios para el formato de datos y los rangos de valores de entrada. * Clasificación de objetos según el tipo de datos. * Desarrollo de un sistema de procesamiento de datos que incluya métodos, criterios, formación y procedimientos de inspección. * Gestión de cambios en los criterios de procesamiento de datos. | Guía del proceso de construcción (Procesamiento) | * Presentación de una guía que define los indicadores de calidad, formato de anotación, datos a procesar, procedimientos, criterios, métodos, herramientas y procesos de inspección en la fase de procesamiento de datos. |
| Revisión de herramientas de etiquetado y entorno de almacenamiento | * Adquisición de herramientas de edición para el procesamiento de datos y definición de su método de uso. * Gestión de cambios en los criterios de procesamiento de datos. * Establecimiento del entorno de almacenamiento para los datos etiquetados. | Herramientas de edición y manual de uso | * Las herramientas de edición para el procesamiento de datos son utilizadas para asignar valores específicos a los datos fuente antes del entrenamiento. * Manual de uso de las herramientas de etiquetado y especificaciones del entorno de almacenamiento. |
| Inspección de calidad de los datos etiquetados | * Inspección con enfoque en la exactitud (precisión semántica y precisión sintáctica). * Herramientas de inspección y resultados de verificación de calidad de los datos etiquetados. * Implementación de medidas correctivas basadas en los resultados de la inspección de calidad. | Datos etiquetados | * Conjunto de anotaciones que incluye los "valores de referencia" (Ground Truth) asignados a los datos fuente, junto con atributos como el formato de archivo, resolución, descripciones y comentarios. |
| Informe de resultados de gestión de calidad | * Elaboración de los estándares de inspección de calidad, estructura organizativa, procedimientos y resultados de la inspección. |
| Herramientas de inspección de calidad | * Definición de herramientas para verificar la exactitud de los datos etiquetados (precisión sintáctica y precisión semántica). * Las herramientas de inspección de calidad pueden incluir aplicaciones o listas de verificación (como formularios de inspección visual, entre otros). |

< Tabla 7 – Actividades de gestión de calidad y entregables en la fase de procesamiento de datos >

* + 1. Entrenamiento de datos

En la fase de entrenamiento de datos, se utiliza el conjunto de datos de entrenamiento, compuesto por datos fuente y datos etiquetados, para entrenar un modelo de inteligencia artificial previamente definido. Además, se llevan a cabo actividades para mejorar o ajustar el rendimiento del modelo de IA entrenado.



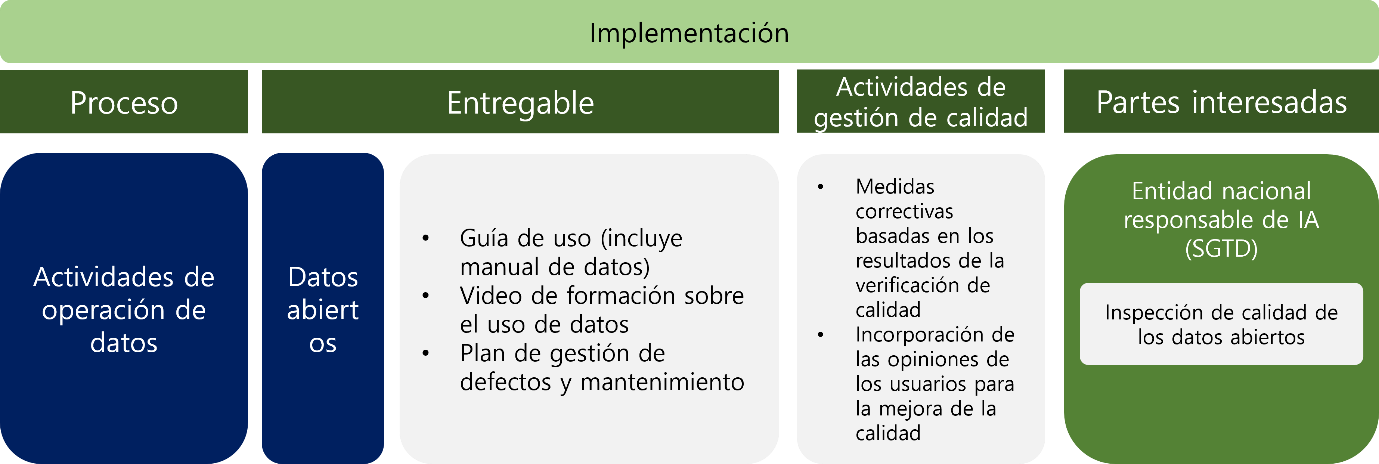
< Figura 10 – Procedimiento de entrenamiento de datos >

|  |  |  |  |
| --- | --- | --- | --- |
| Actividades de gestión de calidad | Contenido principal | Entregables | Definición de entregables |
| Verificación de la compatibilidad entre el propósito de construcción y el modelo de IA. | * Verificación de los estándares de inspección de calidad por área al entrenar los datos en el modelo de inteligencia artificial. * Confirmación de la selección del modelo de IA que se ajusta al propósito de construcción. * Información sobre el entorno establecido para la construcción, operación y prueba del modelo de IA. * Guía sobre la información y licencias del modelo de IA. | Manual del modelo de IA | * Detalles del entorno establecido para la construcción, operación y prueba del modelo de IA. * Configuración del entorno y manual de ejecución para el uso del código fuente del modelo de IA. * Información sobre el modelo, incluyendo detalles de licencia. * Información sobre bibliotecas, versiones de paquetes y entorno de implementación. |
| Revisión de los resultados del entrenamiento y optimización. | * Verificación de si el conjunto de datos de entrenamiento permite alcanzar el nivel de rendimiento objetivo cuando se entrena en el modelo de IA. * Inspección del rendimiento del modelo de IA (algoritmo). * Validación de la efectividad en la fase del modelo. * Análisis de errores e incidencias en la calidad del conjunto de datos de entrenamiento para el desarrollo del modelo de aprendizaje y servicios aplicables, incorporando medidas correctivas. | Código fuente del modelo de IA | * Código de entrenamiento del modelo * Código de prueba de rendimiento del modelo * Otros |
| Archivo del modelo entrenado | * Archivo entrenado para reconocer patrones específicos, que incluye pesos de entrenamiento, información de arquitectura y otros elementos de aprendizaje. |
| Imagen Docker para la implementación del modelo | * Presentación del archivo Docker, que contiene la información del entorno de construcción del modelo de IA, junto con los archivos del modelo como entregables. |
| Informe de resultados de prueba del modelo de IA | * Antes de la verificación de calidad por terceros, la entidad ejecutora del proyecto debe realizar pruebas internas del modelo y aplicar medidas correctivas en caso de deficiencias. |
| Herramientas de inspección de calidad | * Herramienta de inspección de calidad para la validación del conjunto de datos de entrenamiento |
| Implementación de medidas correctivas basadas en los resultados de la verificación de calidad. | * En caso de que los resultados de la verificación de calidad por terceros del conjunto de datos final no cumplan con los estándares, se implementan medidas correctivas y se presenta un informe con los resultados. | Informe de resultados de medidas correctivas | * Solo las entidades que no hayan alcanzado la verificación final de calidad por terceros deben presentar este entregable. * Elaboración del informe de inspección de calidad tras la implementación de mejoras en cada fase, enfocándose en los aspectos señalados como deficientes en el informe final de verificación de calidad. |

<Tabla 8 – Actividades de gestión de calidad y entregables en la fase de entrenamiento de datos>

* 1. Operación y aplicación
     1. Operación y aplicación

La gestión de calidad en la fase de operación y aplicación de los datos de entrenamiento de inteligencia artificial comprende todas las actividades necesarias para mejorar la calidad del conjunto de datos de entrenamiento cargado en la plataforma de aprendizaje de modelos de IA en Perú, con el fin de facilitar su apertura al sector privado.



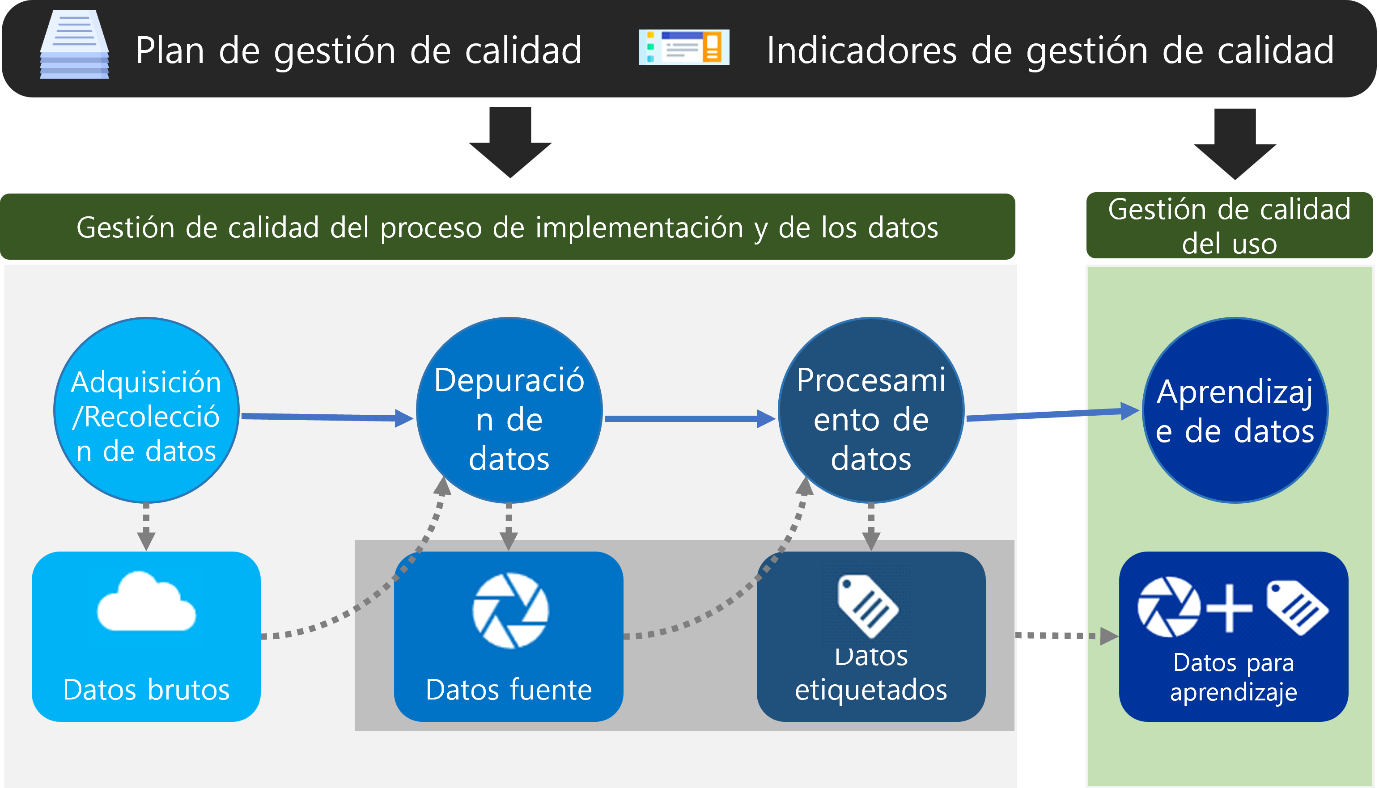
<Figura 11 – Procedimiento de operación y aplicación de datos>

|  |  |  |  |
| --- | --- | --- | --- |
| Actividades de gestión de calidad | Contenido principal | Entregables | Definición de entregables |
| Gestión de defectos y mantenimiento de datos | * Mantenimiento y corrección de defectos en los datos publicados en la plataforma de aprendizaje de modelos de IA en Perú. | Plan de corrección de defectos y mantenimiento | * Plan de corrección de defectos y mantenimiento para el conjunto de datos de entrenamiento de acceso público. |
| Incorporación de sugerencias de mejora de calidad de los usuarios | * Publicación de entregables para apoyar el desarrollo y uso por parte de los usuarios. * Plan de mantenimiento y corrección de defectos en la calidad de los datos. * Incorporación de sugerencias de mejora de calidad presentadas por los usuarios durante el mantenimiento. | Guía de uso (incluyendo documentación de datos) | * Publicación en la plataforma de aprendizaje de modelos de IA en Perú, asegurando el acceso para todos, mediante una documentación de datos y una guía de uso. |
| Video educativo sobre el uso de datos | * Producción de un video educativo de guía para el uso de datos por cada tipo de conjunto de datos y su publicación en la plataforma de aprendizaje de modelos de IA en Perú. |

<Tabla 9 – Actividades de gestión de calidad y entregables en la fase de operación y aplicación de datos>

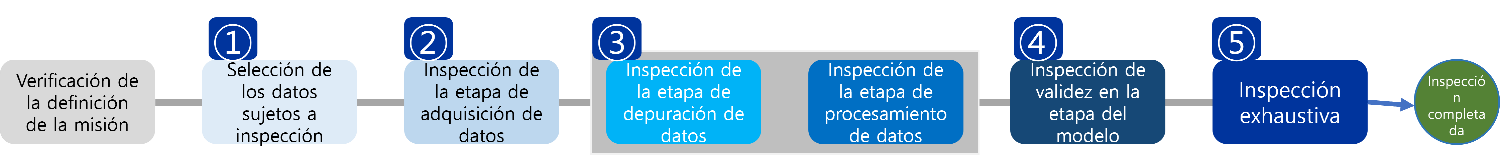
* + - * 1. Estrategia de gestión de la calidad de los datos de inteligencia artificial

Las instituciones deben llevar a cabo la gestión de calidad del proceso de construcción y de los datos en la fase de construcción, así como la gestión de calidad en la aplicación en la fase de operación y uso, conforme al plan de gestión de calidad establecido en la fase de planificación y a los indicadores de gestión de calidad acordados con la entidad de verificación de calidad, según lo siguiente.



<Figura 12 – Estrategia de gestión de la calidad de los datos de inteligencia artificial>

* 1. Estrategia de gestión de la calidad del proceso de construcción y de los datos

La inspección de calidad en la fase de construcción se aplica a todas las áreas del proceso de construcción, excepto el control de calidad final. Dado que los errores de calidad en las etapas posteriores pueden derivarse de problemas en las etapas previas, es fundamental realizar inspecciones de calidad continuas desde la fase de adquisición/recopilación de datos. Esto permite minimizar la necesidad de reprocesamiento y garantizar la calidad en cada fase del proceso.

<Figura 13 – Procedimiento de gestión de la calidad del proceso de construcción y de los datos>

* + 1. Selección de los datos sujetos a inspección

Desde las primeras etapas de la construcción de datos, se debe considerar la inspección de calidad durante la recopilación de datos. Además, es fundamental ajustar adecuadamente la proporción de datos según cada sistema de clasificación para establecer un plan que garantice la mejora de la calidad global del conjunto de datos.

|  |  |
| --- | --- |
| Clasificación | Contenido |
| Nombre del indicador de inspección | No aplica |
| Contenido de inspección | * Selección de estándares y volúmenes adecuados según los indicadores de calidad de cada proceso. * Inclusión de todas las categorías dentro del sistema de clasificación de los datos a construir. * Planificación desde la fase inicial de construcción de datos para garantizar que se consideren las inspecciones de calidad y que los datos de cada categoría se adquieran y recopilen en proporciones adecuadas. |
| Ejemplo de inspección | * En la construcción de datos para el aprendizaje de identificación de hierbas medicinales, si se seleccionan ginseng, ginseng rojo, angélica y olmo, se debe verificar que todos estos datos estén incluidos en la selección de datos sujetos a inspección. |

<Tabla 10 – Contenido y ejemplos de selección de datos sujetos a inspección>

* + 1. Inspección en la fase de adquisición de datos

Durante la fase de adquisición/recopilación de datos, se realiza una inspección de conformidad para redefinir la misión y los requisitos del modelo de aprendizaje. Además, los errores detectados se gestionan de manera sistemática, lo que permite modificar y volver a ejecutar el plan de adquisición/recopilación de datos según sea necesario.

|  |  |
| --- | --- |
| Clasificación | Contenido |
| Nombre del indicador de inspección | Nombre del indicador de inspección: Conformidad (Conformidad con estándares, Conformidad técnica, Diversidad estadística) |
| Contenido de inspección | * Inspección centrada en la conformidad con estándares, conformidad técnica y diversidad estadística. * Para realizar mejoras basadas en los resultados de la inspección en la fase de adquisición/recopilación, es necesario redefinir la misión y los requisitos de construcción de datos según el modelo de aprendizaje, modificar el plan de construcción y repetir el proceso de adquisición/recopilación de datos. * Los errores detectados deben documentarse con sus características y ejemplos para priorizarlos en inspecciones futuras y minimizar su recurrencia. |
| Ejemplo de inspección | * Verificación de la conformidad en el uso de imágenes capturadas o de imágenes abiertas (captura de imágenes, diversidad, cumplimiento legal). * Verificación de la conformidad del ítem, conformidad de la imagen y posible vulneración de datos personales. |

<Tabla 11 – Contenido y ejemplos de inspección en la fase de adquisición/recopilación de datos>

* + 1. Inspección en la fase de depuración y procesamiento de datos

Para garantizar un entrenamiento preciso del modelo de inteligencia artificial, la inspección se centra en la exactitud semántica y sintáctica. Se requiere una verificación de calidad periódica para minimizar las diferencias de calidad entre los operadores. Además, dado que los errores detectados pueden originarse no solo en la fase actual, sino también en fases anteriores, es fundamental mantener una gestión de calidad continua a lo largo de todo el proceso.

|  |  |
| --- | --- |
| Clasificación | Contenido |
| Nombre del indicador de inspección | Nombre del indicador de inspección: Exactitud (Exactitud semántica, Exactitud sintáctica) |
| Contenido de inspección | * Inspección enfocada en la exactitud (exactitud semántica y exactitud sintáctica). * La exactitud es un elemento esencial para que el modelo de inteligencia artificial pueda aprender correctamente, y debido a las características del proceso de construcción, pueden existir diferencias de calidad entre los operadores. * Se deben realizar inspecciones de calidad periódicas o aleatorias de manera continua por parte del responsable de gestión de calidad. * Los errores detectados en esta fase pueden deberse a fallos en el proceso actual o pueden ser consecuencia de errores originados en fases anteriores que afectan a la fase actual. |
| Ejemplo de inspección | * Se debe analizar el impacto de las fases anteriores y, bajo la supervisión del equipo de gestión de calidad o del equipo de gestión del proyecto, se deben reajustar y volver a ejecutar los planes de proceso para minimizar errores. |

<Tabla 12 – Contenido y ejemplos de inspección en la fase de depuración y procesamiento de datos>

* + 1. Inspección en la fase del modelo

Para evaluar los resultados del entrenamiento del modelo de inteligencia artificial, se lleva a cabo una evaluación del rendimiento del modelo en diversas áreas, como clasificación, detección, reconocimiento, preguntas y respuestas, traducción automática, resumen de documentos, imágenes, videos, texto, voz y sensores.

|  |  |
| --- | --- |
| Clasificación | Contenido |
| Nombre del indicador de inspección | Validez (División del conjunto de datos completo en datos de entrenamiento, validación y evaluación; medición del rendimiento con el conjunto de datos de validación o ajuste fino después del entrenamiento). |
| Contenido de inspección | * Verificación de los estándares de inspección de calidad en distintas áreas, incluyendo clasificación, detección, reconocimiento, preguntas y respuestas, traducción automática, resumen de documentos, imágenes, videos, texto, voz y sensores, al entrenar los datos en el modelo de inteligencia artificial. |
| Ejemplo de inspección | * Inspección para comprobar si el modelo de IA entrenado con los datos de entrenamiento alcanza el nivel de rendimiento esperado, utilizando métricas como IoU, AUC, precisión media, Fβ-score, precisión por píxel, precisión media, concordancia y similitud de oraciones, entre otras. |

<Tabla 13 – Contenido y ejemplos de inspección en la fase del modelo>

* + 1. Inspección exhaustiva

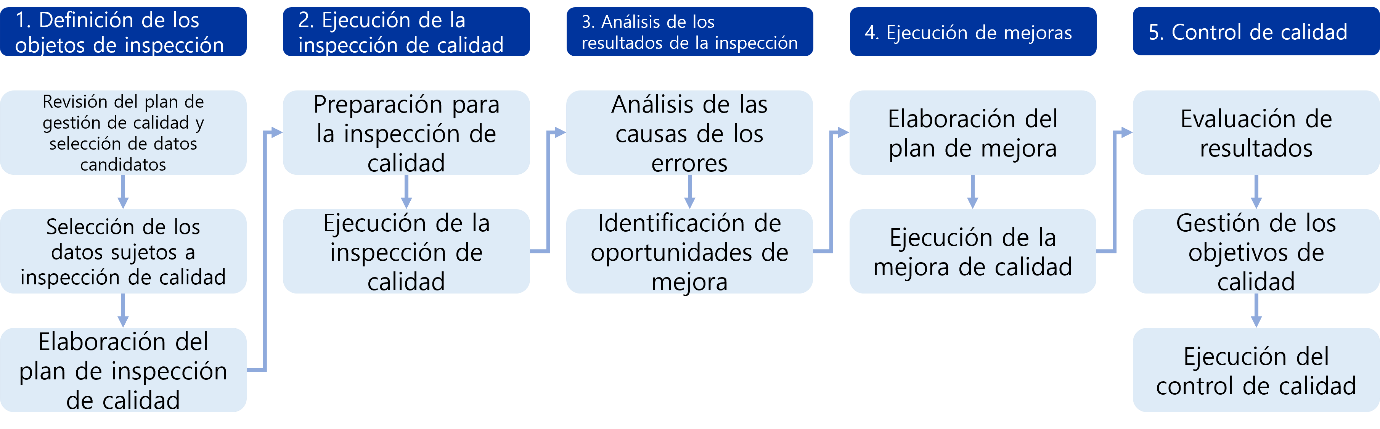
En el proceso de verificación de la calidad de los datos, diversas entidades participan en la inspección bajo estándares estrictos. Con base en los resultados obtenidos, se llevan a cabo mejoras en la calidad de los datos.

|  |  |
| --- | --- |
| Clasificación | Contenido |
| Nombre del indicador de inspección | La inspección se lleva a cabo siguiendo los criterios de verificación de la depuración de datos. |
| Contenido de inspección | * Verificación de la distribución de los datos que no cumplen con los estándares. * Inspección de la uniformidad de los indicadores y criterios de verificación realizada por evaluadores externos, expertos en el dominio y solicitantes de datos. |
| Ejemplo de inspección | * Ejecución del proceso de inspección mediante crowdsourcing, cumpliendo con los criterios de verificación de la depuración de datos. * Prueba de exactitud: La precisión de la concordancia del formato de anotación debe ser superior al 99%. |

<Tabla 14 – Contenido y ejemplos de inspección exhaustiva>

* 1. Estrategia de gestión de calidad en la aplicación

La inspección de calidad en el uso de los datos de entrenamiento de inteligencia artificial se lleva a cabo en cinco fases: definición de los datos sujetos a inspección, ejecución de la inspección de calidad, análisis de resultados, implementación de mejoras y control de calidad. Estas fases se desarrollan según el siguiente procedimiento.



<Figura 14 – Procedimiento detallado de inspección de calidad en la fase de operación y aplicación>

* + 1. Definición de los datos sujetos a inspección

El responsable de gestión de calidad revisa el plan de gestión de calidad del proyecto de construcción, coordina con los encargados pertinentes para incorporar los cambios necesarios y, considerando la importancia y urgencia de la inspección, selecciona los datos sujetos a inspección. Con base en esta selección, se elabora el plan de inspección de calidad.

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Revisión del plan de gestión de calidad y selección de datos candidatos | El responsable de gestión de calidad recopila los documentos relacionados con el plan de gestión de calidad del proyecto de construcción y coordina con los encargados de la construcción de datos y del modelo de aprendizaje de inteligencia artificial para obtener información adicional sobre requisitos actualizados, procedimientos de trabajo y métodos de operación. |
| Selección de los datos sujetos a inspección de calidad | Entre los datos identificados como candidatos en la revisión del plan de gestión de calidad, se seleccionan los datos sujetos a inspección según su importancia y urgencia en términos de gestión de calidad. |
| Elaboración del plan de inspección de calidad | Una vez finalizada la selección de los datos sujetos a inspección, se procede a la elaboración del plan de inspección de calidad para los datos seleccionados. |

<Tabla 15 – Procedimiento de definición de los datos sujetos a inspección>

* + 1. Ejecución de la inspección de calidad

La fase de ejecución de la inspección de calidad implica la asignación de personal especializado, el establecimiento de un sistema de cooperación, la preparación de los datos y herramientas de inspección, y la habilitación de un entorno adecuado para la inspección con el fin de garantizar una ejecución eficiente del proceso.

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Preparación de la inspección de calidad | La preparación de la inspección de calidad es una etapa previa a la ejecución de la inspección, en la que la entidad ejecutora del proyecto establece el entorno necesario para llevar a cabo la inspección de manera eficiente.   * Asignación de personal especializado en inspección de calidad * Establecimiento de un sistema de cooperación * Preparación de los datos sujetos a inspección * Preparación de las herramientas de inspección * Habilitación del lugar de ejecución |
| Ejecución de la inspección de calidad |

<Tabla 16 – Procedimiento de ejecución de la inspección de calidad>

* + 1. Análisis de los resultados de la inspección

Proceso en el que se analizan los resultados de la inspección de calidad para identificar las causas de los errores y, con base en este análisis, se establecen estrategias de mejora para optimizar el sistema.

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Análisis de las causas de los errores | Para rastrear las causas de los errores, se debe verificar si los errores identificados previamente en la fase de inspección de calidad son realmente errores. Incluso si se confirman como tales, es necesario considerar la misión y las características de los datos para determinar si deben incluirse en el análisis. |
| Identificación de oportunidades de mejora | Las oportunidades de mejora identificadas deben vincularse directamente con las acciones concretas a implementarse en la fase de "Ejecución de mejoras (Implement)". |

<Tabla 17 – Procedimiento de análisis de los resultados de la inspección>

* + 1. Implementación de mejoras

A partir del análisis de los resultados de la inspección de calidad, se identifican las áreas que requieren mejoras y se establecen tareas específicas de mejora, priorizándolas según su importancia. Posteriormente, se crea un entorno adecuado para la implementación y se llevan a cabo las actividades de mejora, comenzando por las tareas de alta prioridad y corto plazo.

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de mejora | Se recopilan las oportunidades de mejora identificadas en la fase de análisis de los resultados de la inspección para definir tareas de mejora y se elabora un plan detallado de implementación para cada tarea categorizada. |
| Ejecución de la mejora de calidad | De las mejoras planificadas, se priorizan aquellas de corto plazo, se prepara el entorno necesario para su ejecución y se llevan a cabo las acciones de mejora correspondientes. |

<Tabla 18 – Procedimiento de implementación de mejoras>

* + 1. Control de calidad

Se establecen mecanismos de gestión de calidad, se optimiza el plan de operación de datos y se ajustan los datos según los objetivos de calidad. Además, se lleva a cabo un monitoreo continuo y se implementan diversas actividades de mejora para actualizar la guía del proceso de construcción y mejorar la calidad real de los datos.

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Evaluación de resultados | Se divide en establecimiento del sistema de gestión de calidad, planificación de la operación de datos y ajuste de datos.  Se representa de manera integral la revisión de la guía del proceso de construcción y las áreas de mejora reales, evaluando cuantitativamente los efectos de las mejoras implementadas. |
| Gestión de objetivos de calidad | Se establecen objetivos de calidad para los elementos sujetos a control de calidad.  Se definen estrategias para mantener o mejorar la calidad de los datos. |
| Ejecución del control de calidad | Se lleva a cabo un monitoreo continuo de los elementos bajo gestión de calidad.  Se fomenta el consenso sobre la gestión de los objetivos de calidad a través de formación y actividades de divulgación dirigidas a los miembros de la entidad ejecutora del proyecto.  Se promueve una participación activa y una comunicación efectiva para fortalecer el compromiso con la gestión de calidad. |

<Tabla 19 – Procedimiento de control de calidad>

* + - * 1. Indicadores de gestión de calidad

Para verificar la calidad de los datos de acuerdo con el propósito de su construcción, es esencial implementar una gestión de calidad sistemática a lo largo de todo el ciclo de vida de la construcción de datos. Para ello, se deben definir indicadores de gestión de calidad y evaluar la calidad de los datos en cuatro aspectos principales: preparación, integridad, conformidad y adecuación, con el fin de medir el nivel de calidad de los datos. Los indicadores de gestión de calidad constituyen un criterio clave para garantizar que los datos cumplan con las necesidades de los usuarios y sean adecuados para el aprendizaje de inteligencia artificial. El método de gestión de calidad de los datos se basa en la realización de inspecciones cuantitativas según los indicadores específicos y los elementos sujetos a inspección. Además, se establecen relaciones entre los indicadores y los elementos inspeccionados para aplicar los métodos de inspección adecuados a cada indicador.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Indicadores de calidad | | | Inspección de calidad según el ciclo de vida de los datos | | | | | |
| Definición de la misión | Elaboración del plan de construcción | Adquisición/recopilación de datos | Depuración de datos | Procesamiento de datos | Entrenamiento de datos |
| Preparación | Estructuración de la planificación | Preparación de procedimientos | ● | ● | ● | ● | ● |  |
| Preparación organizativa | ● | ● | ● | ● | ● |  |
| Preparación de herramientas |  | ● | ● | ● | ● |  |
| Gestión de riesgos | ● | ● | ● | ● | ● |  |
| Cumplimiento del sistema | Cumplimiento de seguridad | ● | ● | ● | ● | ● |  |
| Cumplimiento legal y normativo | ● | ● | ● | ● | ● |  |
| Integridad | Integridad en la recopilación | |  |  | ● |  |  |  |
| Integridad en la depuración | |  |  |  | ● |  |  |
| Integridad en el procesamiento | |  |  |  |  | ● |  |
| Utilidad | Facilidad de uso | | ● |  |  |  |  | ● |
| Flexibilidad | | ● |  |  |  |  | ● |
| Conformidad | Conformidad con los estándares | Diversidad |  | ● | ● |  |  |  |
| Fiabilidad |  | ● | ● |  |  |  |
| Suficiencia |  | ● | ● |  |  |  |
| Uniformidad |  | ● | ● |  |  |  |
| Veracidad |  | ● | ● |  |  |  |
| Equidad |  | ● | ● |  |  |  |

<Tabla 20 – Elementos sujetos a inspección y métodos de verificación por indicador de calidad>

* 1. Preparación

Indicador que verifica si se han establecido y se gestionan adecuadamente las políticas, regulaciones (incluyendo revisiones sobre derechos de autor, derechos de imagen, protección de datos personales y seguridad de la información), estructuras organizativas y procedimientos esenciales para la gestión de la calidad de los datos de entrenamiento de inteligencia artificial. Además, evalúa si estos elementos se mantienen actualizados y correctamente administrados.

* + 1. Estructuración de la planificación – Preparación de procedimientos
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si los procedimientos para la construcción de datos de entrenamiento de inteligencia artificial están establecidos de manera sistemática y se gestionan y ejecutan correctamente. |
| Elemento sujeto a inspección | Entregables | * Documentos de referencia: Plan de ejecución del proyecto (Plan de construcción, Plan de gestión de calidad), entre otros. |
| Proceso | * Ámbitos de inspección: Definición de la misión, planificación de la construcción, adquisición/recopilación de datos, depuración de datos, procesamiento de datos. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 21 – Criterios y métodos de inspección para la preparación de procedimientos en la estructuración de la planificación>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Definición de la misión | * ¿Se han recopilado los requisitos del organismo contratante (demandante)? * ¿Se han analizado los requisitos del organismo contratante (demandante)? * ¿Se ha realizado una investigación previa sobre estudios similares o equivalentes relacionados con el proyecto? * ¿Se ha establecido una estrategia para la selección del modelo de IA más adecuado para el proyecto? * ¿Se han presentado los resultados de la selección del modelo de IA adecuado para el proyecto? * ¿Se han identificado las restricciones y limitaciones en la selección del modelo de IA adecuado para el proyecto? * ¿Se han definido los indicadores de rendimiento y objetivos del modelo de IA? * ¿Se han establecido métricas e indicadores para garantizar el rendimiento del modelo de IA, considerando sus características y estudios previos? * ¿Se han proporcionado evidencias objetivas para respaldar los indicadores y objetivos de rendimiento del modelo de IA? (por ejemplo, asesoramiento de expertos). |
| Elaboración del plan de construcción | * ¿Se ha definido el conjunto de datos de entrenamiento de inteligencia artificial? * ¿Se ha establecido la proporción de los conjuntos de datos de entrenamiento, validación y prueba? (Ejemplo: 8:1:1 u otra proporción) * ¿Se ha presentado un plan de gestión para garantizar que los datos de prueba del conjunto de datos de entrenamiento de IA no sean utilizados en el entrenamiento? (Reglamentos de gestión y operación, entorno, control de acceso, etc.) * ¿Se ha establecido un plan de configuración del entorno para la construcción de datos en bruto, datos fuente, datos etiquetados y modelos de entrenamiento de IA? * ¿Se ha definido el sistema de clasificación de datos de entrenamiento de IA? * ¿Se ha planificado la revisión del modelo de IA utilizando un conjunto de datos inicial (7-10% del total) para evaluar su idoneidad? * ¿Se ha establecido un plan de adquisición/recopilación, depuración y procesamiento de datos según la clasificación definida? * ¿Se han determinado los criterios y métodos de selección del conjunto de datos inicial (7-10%) para la revisión del modelo de IA? * ¿Se ha definido un procedimiento para reflejar los resultados de la revisión del modelo de IA utilizando el conjunto de datos inicial (7-10%)? |
| Adquisición/recopilación de datos | * ¿Se ha establecido un procedimiento de revisión legal y normativa para la selección de datos en relación con protección de datos personales, derechos de imagen, derechos de autor, difamación y seguridad? * ¿Se ha definido un procedimiento de revisión sobre las tecnologías y métodos de recopilación de datos al momento de la selección? * ¿Se ha establecido un procedimiento de revisión legal y normativa para la adquisición/recopilación de datos, considerando aspectos como protección de datos personales, derechos de imagen, derechos de autor, difamación y seguridad? * ¿Se han definido los métodos y criterios para la adquisición/recopilación de datos? * ¿Se ha desarrollado una estrategia para la obtención de datos faltantes en la fase de adquisición/recopilación? * ¿Se ha establecido un procedimiento para gestionar cambios en los criterios de adquisición/recopilación de datos? * ¿Se ha diseñado un plan de formación y capacitación sobre los métodos de adquisición/recopilación de datos? * ¿Se ha definido un procedimiento de verificación para evaluar los resultados de la adquisición/recopilación de datos? * ¿Se ha planificado una formación y capacitación sobre los procedimientos de verificación de los resultados de adquisición/recopilación de datos? * ¿Se ha establecido un procedimiento para reflejar en el proceso los resultados de la revisión del modelo de IA utilizando un conjunto de datos inicial (7-10%) para evaluar su idoneidad? |
| Depuración de datos | * ¿Se han establecido criterios y procedimientos para la anonimización de datos, garantizando la protección de datos personales durante la depuración? * ¿Se han definido los métodos y criterios para la depuración de datos? * ¿Se ha establecido un procedimiento para gestionar cambios en los criterios de depuración de datos? * ¿Se ha diseñado un plan de formación y capacitación sobre los métodos de depuración de datos? * ¿Se han determinado los procedimientos y criterios de verificación para evaluar los resultados de la depuración de datos? * ¿Se ha establecido un procedimiento para gestionar cambios en los criterios de verificación de los resultados de depuración? * ¿Se ha planificado una formación y capacitación sobre los procedimientos de verificación de los resultados de depuración de datos? |
| Procesamiento de datos | * ¿Se han definido los métodos y criterios para el etiquetado de datos? * ¿Se ha establecido un procedimiento para gestionar cambios en los criterios de etiquetado de datos? * ¿Se ha diseñado un plan de formación y capacitación sobre los métodos de etiquetado de datos? * ¿Se han determinado los procedimientos de verificación para evaluar los resultados del etiquetado de datos? * ¿Se ha establecido un procedimiento para gestionar cambios en los criterios de verificación de los resultados de etiquetado de datos? * ¿Se ha planificado una formación y capacitación sobre los procedimientos de verificación de los resultados de etiquetado de datos? * ¿Se ha desarrollado un método de muestreo para verificar si los datos etiquetados son adecuados para el modelo de aprendizaje de inteligencia artificial? |

<Tabla 22 – Contenido de inspección para la preparación de procedimientos en la estructuración de la planificación>

* + 1. Estructuración de la planificación – Preparación organizativa
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si la estructura organizativa, los roles y las responsabilidades para la construcción de datos de entrenamiento de inteligencia artificial están establecidos de manera sistemática y se gestionan y operan correctamente. |
| Elemento sujeto a inspección | Entregables | * Documentos de referencia: Plan de ejecución del proyecto (Plan de construcción, Plan de gestión de calidad), entre otros. |
| Proceso | * Ámbitos de inspección: Elaboración del plan de construcción, adquisición/recopilación de datos, depuración de datos, procesamiento de datos. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 23 – Criterios y métodos de inspección para la preparación organizativa en la estructuración de la planificación>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Adquisición/recopilación de datos | * ¿Se ha establecido un equipo especializado para la revisión legal y normativa en la adquisición/recopilación de datos? * ¿Se ha conformado un equipo de adquisición/recopilación de datos y se han designado los responsables correspondientes? * ¿Se ha elaborado un plan de gestión del personal para la adquisición/recopilación de datos? * ¿Se han definido los roles y responsabilidades del equipo encargado de la adquisición/recopilación de datos? * ¿Se ha conformado un equipo de formación y capacitación sobre los métodos de adquisición/recopilación de datos y se han designado los responsables? * ¿Se han definido los roles y responsabilidades del equipo encargado de la formación y capacitación en adquisición/recopilación de datos? * ¿Se han establecido los roles y responsabilidades del equipo encargado de la verificación de los resultados de la adquisición/recopilación de datos? |
| Depuración de datos | * ¿Se ha establecido un equipo especializado para la definición de criterios de anonimización y protección de datos personales en la fase de depuración de datos? * ¿Se ha conformado un equipo de depuración de datos y se han designado los responsables correspondientes? * ¿Se ha elaborado un plan de gestión del personal para la depuración de datos? * ¿Se han definido los roles y responsabilidades del equipo encargado de la depuración de datos? * ¿Se ha conformado un equipo de formación y capacitación sobre los métodos de depuración de datos y se han designado los responsables? * ¿Se han definido los roles y responsabilidades del equipo encargado de la formación y capacitación en depuración de datos? * ¿Se han establecido los roles y responsabilidades del equipo encargado de la verificación de los resultados de la depuración de datos? |
| Procesamiento de datos | * ¿Se ha conformado un equipo de etiquetado de datos y se han designado los responsables correspondientes? * ¿Se ha elaborado un plan de gestión del personal para la fase de etiquetado de datos? * ¿Se han definido los roles y responsabilidades del equipo encargado del etiquetado de datos? * ¿Se ha conformado un equipo de formación y capacitación sobre los métodos de etiquetado de datos y se han designado los responsables? * ¿Se han definido los roles y responsabilidades del equipo encargado de la formación y capacitación en etiquetado de datos? * ¿Se han establecido los roles y responsabilidades del equipo encargado de la verificación de los resultados del etiquetado de datos? |

<Tabla 24 – Contenido de inspección para la preparación organizativa en la estructuración de la planificación>

* + 1. Estructuración de la planificación – Preparación de herramientas
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si se ha establecido un plan de configuración y gestión de herramientas y entornos para la construcción de datos de entrenamiento de inteligencia artificial y si estos se administran y operan correctamente. |
| Elemento sujeto a inspección | Entregables | * Documentos de referencia: Plan de ejecución del proyecto (Plan de construcción, Plan de gestión de calidad), entre otros. |
| Proceso | * Ámbitos de inspección: Elaboración del plan de construcción, adquisición/recopilación de datos, depuración de datos, procesamiento de datos. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 25 – Criterios y métodos de inspección para la preparación de herramientas en la estructuración de la planificación>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de mejora | * ¿Se ha definido el modelo de inteligencia artificial seleccionado para el entrenamiento? * ¿Se han propuesto modelos de inteligencia artificial alternativos en función de posibles cambios en los requisitos? |
| Adquisición/recopilación de datos | * ¿Se han definido las herramientas de trabajo para la adquisición/recopilación de datos? * ¿Existe un manual de usuario/administrador para el uso de las herramientas de adquisición/recopilación de datos? * ¿Se ha elaborado un plan de formación y capacitación sobre el uso de las herramientas de adquisición/recopilación de datos? * ¿Se ha llevado a cabo la formación y capacitación en el uso de las herramientas de adquisición/recopilación de datos? * ¿Se ha realizado una revisión legal y normativa (incluyendo derechos de autor) sobre el uso de las herramientas de adquisición/recopilación de datos? |
| Depuración de datos | * ¿Se han definido las herramientas de trabajo para la depuración de datos? * ¿Existe un manual de usuario/administrador para el uso de las herramientas de depuración de datos? * ¿Se ha elaborado un plan de formación y capacitación sobre el uso de las herramientas de depuración de datos? * ¿Se ha llevado a cabo la formación y capacitación en el uso de las herramientas de depuración de datos? * ¿Se ha realizado una revisión legal y normativa (incluyendo derechos de autor) sobre el uso de las herramientas de depuración de datos? |
| Procesamiento de datos | * ¿Se han definido las herramientas de trabajo para el etiquetado de datos? * ¿Existe un manual de usuario/administrador para el uso de las herramientas de etiquetado de datos? * ¿Se ha elaborado un plan de formación y capacitación sobre el uso de las herramientas de etiquetado de datos? * ¿Se ha llevado a cabo la formación y capacitación en el uso de las herramientas de etiquetado de datos? * ¿Se ha realizado una revisión legal y normativa (incluyendo derechos de autor) sobre el uso de las herramientas de etiquetado de datos? |

<Tabla 26 – Contenido de inspección para la preparación de herramientas en la estructuración de la planificación>

* + 1. Estructuración de la planificación – Gestión de riesgos
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si se ha establecido un sistema de gestión de riesgos para la construcción de datos de entrenamiento de inteligencia artificial y si este se gestiona y ejecuta adecuadamente. |
| Elemento sujeto a inspección | Entregables | * Documentos de referencia: Plan de ejecución del proyecto (Plan de construcción, Plan de gestión de calidad), entre otros. |
| Proceso | * Ámbitos de inspección: Elaboración del plan de construcción, adquisición/recopilación de datos, depuración de datos, procesamiento de datos. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 27 – Criterios y métodos de inspección para la gestión de riesgos en la estructuración de la planificación>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de construcción | * ¿Se ha establecido un plan de gestión de riesgos para el proyecto? * ¿Se han identificado los factores de riesgo del proyecto? * ¿Se están llevando a cabo acciones para mitigar los riesgos identificados? |
| Adquisición/recopilación de datos | * ¿Se ha establecido un plan de gestión de riesgos para la adquisición/recopilación de datos? * ¿Se han identificado los factores de riesgo en la adquisición/recopilación de datos? * ¿Se están llevando a cabo acciones para mitigar los riesgos identificados en la adquisición/recopilación de datos? |
| Depuración de datos | * ¿Se ha establecido un plan de gestión de riesgos para la depuración de datos? * ¿Se han identificado los factores de riesgo en la depuración de datos? * ¿Se están llevando a cabo acciones para mitigar los riesgos identificados en la depuración de datos? |
| Procesamiento de datos | * ¿Se ha establecido un plan de gestión de riesgos para el etiquetado de datos? * ¿Se han identificado los factores de riesgo en el proceso de etiquetado de datos? * ¿Se están llevando a cabo acciones para mitigar los riesgos identificados en el etiquetado de datos? |

<Tabla 28 – Contenido de inspección para la gestión de riesgos en la estructuración de la planificación>

* + 1. Cumplimiento del sistema – Cumplimiento de seguridad
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si se han establecido y se gestionan adecuadamente procedimientos y normativas para la protección de información sensible, incluyendo la protección de datos personales, medidas de anonimización y gestión de seguridad de los datos de entrenamiento de inteligencia artificial. |
| Elemento sujeto a inspección | Entregables | * Documentos de referencia: Plan de ejecución del proyecto (Plan de construcción, Plan de gestión de calidad), entre otros. |
| Proceso | * Ámbitos de inspección: Elaboración del plan de construcción, adquisición/recopilación de datos, depuración de datos, procesamiento de datos. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 29 – Criterios y métodos de inspección para el cumplimiento de seguridad en el cumplimiento del sistema>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de construcción | * ¿Se ha establecido un sistema de gestión de seguridad para los conjuntos de datos de entrenamiento de inteligencia artificial (entrenamiento, validación y prueba), incluyendo reglamentos de gestión y operación, entorno y control de acceso? * ¿Se ha implementado un sistema de protección de información sensible, incluyendo reglamentos de gestión y operación, entorno y control de acceso?   ※ Cumplimiento de la legislación aplicable, normativas y recomendaciones en materia de protección de datos personales y seguridad. |
| Adquisición/recopilación de datos | * ¿Se está llevando a cabo la gestión de seguridad en la adquisición/recopilación de datos? (Designación de responsables, configuración del entorno operativo, gestión de accesos, etc.) * ¿Se están implementando medidas para la protección de información sensible? (Designación de responsables, gestión de registros, etc.)   ※ Cumplimiento de la legislación aplicable, normativas y recomendaciones en materia de protección de datos personales y seguridad. |
| Depuración de datos | * ¿Se está llevando a cabo la anonimización y protección de datos personales durante la depuración de datos? * ¿Se está implementando la gestión de seguridad en la depuración de datos? (Designación de responsables, configuración del entorno operativo, gestión de accesos, etc.)   ※ Cumplimiento de la legislación aplicable, normativas y recomendaciones en materia de protección de datos personales y seguridad. |
| Procesamiento de datos | * ¿Se está implementando la gestión de seguridad en el etiquetado de datos? (Designación de responsables, configuración del entorno operativo, gestión de accesos, etc.)   ※ Cumplimiento de la legislación aplicable, normativas y recomendaciones en materia de protección de datos personales y seguridad. |

<Tabla 30 – Contenido de inspección para el cumplimiento de seguridad en el cumplimiento del sistema>

* + 1. Cumplimiento del sistema – Cumplimiento legal y normativo
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si, durante la construcción de datos de entrenamiento de inteligencia artificial, se han realizado revisiones adecuadas de las leyes y normativas aplicables (protección de datos personales, derechos de imagen, derechos de autor, etc.) y se han propuesto soluciones para su cumplimiento. |
| Elemento sujeto a inspección | Entregables | * Documentos de referencia: Plan de ejecución del proyecto (Plan de construcción, Plan de gestión de calidad), entre otros. |
| Proceso | * Ámbitos de inspección: Elaboración del plan de construcción, adquisición/recopilación de datos, depuración de datos, procesamiento de datos. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 31 – Criterios y métodos de inspección para el cumplimiento legal y normativo en el cumplimiento del sistema>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de construcción | * ¿Se han establecido procedimientos y soluciones para la revisión legal y normativa en la construcción de datos de entrenamiento de inteligencia artificial? * ¿Se ha implementado un procedimiento de consentimiento para el uso de datos personales en la construcción de datos de entrenamiento de inteligencia artificial? * ¿Se ha establecido y ejecutado un procedimiento de consentimiento para el uso de derechos de imagen en la construcción de datos de entrenamiento de inteligencia artificial? * ¿Se ha desarrollado y aplicado un procedimiento para evaluar la posibilidad de difamación en la construcción de datos de entrenamiento de inteligencia artificial?   ※ Cumplimiento de la legislación aplicable en protección de datos personales, seguridad, derechos de autor y derechos de imagen. |
| Adquisición/recopilación de datos | * ¿Se han establecido procedimientos y soluciones para la revisión legal y normativa en la adquisición/recopilación de datos? * En caso de que los datos adquiridos/recopilados estén protegidos por derechos de autor, ¿se han desarrollado estrategias para su recopilación dentro de los límites legales? * ¿Se ha implementado y ejecutado un procedimiento de consentimiento para el uso de obras protegidas por derechos de autor en la adquisición/recopilación de datos? * ¿Se ha establecido y aplicado un procedimiento contractual para el uso de obras protegidas por derechos de autor en la adquisición/recopilación de datos?   ※ Cumplimiento de la legislación aplicable en protección de datos personales, seguridad, derechos de autor y derechos de imagen. |
| Depuración de datos | * ¿Se han presentado procedimientos de revisión legal y regulatoria, así como soluciones para la anonimización de información sensible en la limpieza de datos? * ¿Se ha establecido un plan para aplicar técnicas de anonimización en la limpieza de datos?   ※ Cumplimiento de la legislación aplicable en protección de datos personales, seguridad, derechos de autor y derechos de imagen. |
| Procesamiento de datos | * ¿Se han establecido procedimientos y soluciones para la revisión legal y normativa en el etiquetado de datos?   ※ Cumplimiento de la legislación aplicable en protección de datos personales, seguridad, derechos de autor y derechos de imagen. |

<Tabla 32 – Contenido de inspección para el cumplimiento legal y normativo en el cumplimiento del sistema>

* 1. Integridad

Indicador que verifica si, en la construcción de datos de entrenamiento de inteligencia artificial, se ha establecido una estructura física adecuada y si los datos se han diseñado y almacenado conforme al formato de datos definido y a los rangos de valores de entrada especificados.

* + 1. Integridad en la recopilación
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si se ha establecido, gestionado y aplicado un plan para prevenir sesgos en los datos. * Verificación de si se ha desarrollado, gestionado y aplicado un sistema para los métodos, criterios, formación y verificación en la adquisición/recopilación de datos. * Verificación de si se ha implementado, gestionado y aplicado un sistema para gestionar cambios en los criterios de adquisición/recopilación de datos. |
| Elemento sujeto a inspección | Entregables | * Plan de construcción, Plan de gestión de calidad, Guía de construcción, entre otros. |
| Proceso | * Adquisición/recopilación de datos |
| Método de inspección | | * Lista de verificación (checklist) |

<Tabla 33 – Criterios y métodos de inspección para la integridad en la recopilación>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Adquisición/recopilación de datos | * ¿Se ha establecido un plan para prevenir sesgos en los datos? * ¿Se están aplicando los métodos y criterios definidos para la adquisición/recopilación de datos? * ¿Se ha implementado un procedimiento para gestionar cambios en los criterios de adquisición/recopilación de datos? * ¿Se ha llevado a cabo la formación y capacitación sobre los métodos de adquisición/recopilación de datos? * ¿Se está ejecutando la verificación de los resultados de adquisición/recopilación de datos según los procedimientos y criterios establecidos? * En caso de cambios en los criterios de verificación, ¿se han aplicado conforme a los procedimientos definidos? * ¿Se ha realizado la formación y capacitación sobre los procedimientos de verificación de los resultados de adquisición/recopilación de datos? * ¿Se ha llevado a cabo la selección de datos conforme al plan de filtrado de datos iniciales (7-10% del total) para la revisión de la adecuación al modelo de aprendizaje de IA? |

<Tabla 34 – Contenido de inspección para la integridad en la recopilación>

* + 1. Integridad en la depuración
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si se han establecido, gestionado y aplicado procedimientos y normativas para la protección de información sensible, incluyendo medidas de anonimización para la protección de datos personales. * Verificación de si se ha desarrollado, gestionado y aplicado un sistema para los métodos, criterios, formación y verificación en la depuración de datos. * Verificación de si se ha implementado, gestionado y aplicado un sistema para gestionar cambios en los criterios de depuración de datos. |
| Elemento sujeto a inspección | Entregables | * Plan de construcción, Plan de gestión de calidad, Guía de construcción, entre otros. |
| Proceso | * Depuración de datos |
| Método de inspección | | * Lista de verificación (checklist) |

<Tabla 35 – Criterios y métodos de inspección para la integridad en la depuración>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Depuración de datos | * ¿Se ha realizado la anonimización, incluyendo la protección de datos personales, durante la limpieza de datos? * ¿Se están aplicando los métodos y criterios definidos para la depuración de datos? * ¿Se ha establecido un procedimiento para gestionar cambios en los criterios de depuración de datos? * ¿Se ha llevado a cabo la formación y capacitación sobre los métodos de depuración de datos? * ¿Se está ejecutando la verificación de los resultados de depuración de datos según los procedimientos y criterios establecidos? * En caso de cambios en los criterios de verificación, ¿se han aplicado conforme a los procedimientos definidos? * ¿Se ha realizado la formación y capacitación sobre los procedimientos de verificación de los resultados de depuración de datos? * ¿Se ha llevado a cabo la selección de datos conforme al plan de filtrado de datos iniciales (7-10% del total) para la revisión de la adecuación al modelo de aprendizaje de IA? |

<Tabla 36 – Contenido de inspección para la integridad en la depuración>

* + 1. Integridad en el procesamiento
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si se ha establecido, gestionado y aplicado un sistema para los métodos, criterios, formación y verificación en el etiquetado de datos. * Verificación de si se ha implementado, gestionado y aplicado un sistema para gestionar cambios en los criterios de etiquetado de datos. |
| Elemento sujeto a inspección | Entregables | * Plan de construcción, Plan de gestión de calidad, Guía de construcción, entre otros. |
| Proceso | * Procesamiento de datos. |
| Método de inspección | | * Lista de verificación (checklist). |

<Tabla 37 – Criterios y métodos de inspección para la integridad en el procesamiento>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Procesamiento de datos | * ¿Se están aplicando los métodos y criterios definidos para el etiquetado de datos? * ¿Se ha establecido un procedimiento para gestionar cambios en los criterios de etiquetado de datos? * ¿Se ha llevado a cabo la formación y capacitación sobre los métodos de etiquetado de datos? * ¿Se está ejecutando la verificación de los resultados del etiquetado de datos según los procedimientos y criterios establecidos? * En caso de cambios en los criterios de verificación, ¿se han aplicado conforme a los procedimientos definidos? * ¿Se ha realizado la formación y capacitación sobre los procedimientos de verificación de los resultados del etiquetado de datos? * ¿Se ha llevado a cabo la selección de datos conforme al plan de filtrado de datos iniciales (7-10% del total) para la revisión de la adecuación al modelo de aprendizaje de IA? |

<Tabla 38 – Contenido de inspección para la integridad en el procesamiento>

* 1. Utilidad

Indicador que verifica si se han incorporado adecuadamente los requisitos del usuario y si el alcance y nivel de detalle de los datos de entrenamiento de inteligencia artificial cumplen con la definición de la misión y los objetivos del proyecto.

* + 1. Facilidad de uso
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si el modelo de aprendizaje y el conjunto de datos de entrenamiento proporcionados cumplen con los requisitos del usuario. |
| Elemento sujeto a inspección | Entregables | * Guía de construcción. |
| Proceso | * Entrenamiento de datos. |
| Método de inspección | | * Lista de verificación (checklist). |

<Tabla 39 – Criterios y métodos de inspección para la facilidad de uso>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Definición de la misión | * ¿El conjunto de datos es adecuado para el modelo de inteligencia artificial seleccionado, de acuerdo con los resultados acordados con el organismo contratante (usuario demandante)? |
| Entrenamiento de datos | * ¿El modelo de inteligencia artificial cumple con los indicadores de rendimiento establecidos? * ¿El modelo de IA, al utilizar los conjuntos de datos de entrenamiento y validación, cumple con los indicadores de rendimiento y objetivos? * ¿El modelo de IA, al utilizar el conjunto de datos de prueba, cumple con los indicadores de rendimiento y objetivos? (Excepto para la entidad ejecutora) * ¿Se proporciona información básica para que los usuarios puedan buscar y utilizar el conjunto de datos en el modelo de IA? (Formato de aprendizaje, etc.) |

<Tabla 40 – Contenido de inspección para la facilidad de uso>

* + 1. Flexibilidad
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si la selección y aplicación del modelo de aprendizaje permite adaptarse de manera flexible a los requisitos del usuario. |
| Elemento sujeto a inspección | Entregables | * Guía de construcción. |
| Proceso | * Entrenamiento de datos. |
| Método de inspección | | * Lista de verificación (checklist). |

<Tabla 41 – Criterios y métodos de inspección para la flexibilidad>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Definición de la misión | * ¿Se han propuesto alternativas en caso de que el modelo de inteligencia artificial no cumpla con los indicadores de rendimiento? |
| Entrenamiento de datos | * ¿Se están gestionando y registrando los cambios e historial de los datos de entrenamiento de inteligencia artificial? |

<Tabla 42 – Contenido de inspección para la flexibilidad>

* 1. Conformidad con los estándares

Indicador que mide si los datos construidos son adecuados para el propósito de entrenamiento mediante la evaluación de diversidad, fiabilidad, suficiencia, uniformidad, veracidad y equidad.

* + 1. Diversidad
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación del grado de variabilidad en las condiciones ambientales (espacio, tiempo, etc.) y en las características y tipos de objetos. |
| Elemento sujeto a inspección | Entregables | * Guía de construcción, datos en bruto. |
| Tipo de datos | * Tipos de datos: Texto, imagen, video, audio, 3D. |
| Método de inspección | | * Métodos de inspección: Lista de verificación (checklist), inspección exhaustiva. |

<Tabla 43 – Criterios y métodos de inspección para la diversidad>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de construcción | * ¿Se ha establecido un plan para garantizar que los datos reflejen características similares a las del mundo real que la inteligencia artificial debe procesar? (Ejemplo: objetos, personas, lugares, tiempo, entorno, características lingüísticas, etc.) * ¿Se ha diseñado un plan para asegurar que los datos posean una variabilidad similar a la de los datos del mundo real que la inteligencia artificial debe procesar? (Ejemplo: cambios en objetos, personas, lugares, tiempo, entorno, características lingüísticas, etc.) |
| Adquisición/recopilación de datos | * ¿Los datos reflejan características similares a las del mundo real que la inteligencia artificial debe procesar? (Ejemplo: objetos, personas, lugares, tiempo, entorno, características lingüísticas, etc.) * ¿Los datos poseen una variabilidad similar a la de los datos del mundo real que la inteligencia artificial debe procesar? (Ejemplo: cambios en objetos, personas, lugares, tiempo, entorno, características lingüísticas, etc.) |

<Tabla 44 – Contenido de inspección para la diversidad>

* + 1. Fiabilidad
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si los datos han sido recopilados de fuentes confiables, como personas, instituciones o empresas. |
| Elemento sujeto a inspección | Entregables | * Guía de construcción, datos en bruto. |
| Tipo de datos | * Tipos de datos: Texto, imagen, video, audio, 3D. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 45 – Criterios y métodos de inspección para la fiabilidad>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de construcción | * ¿Se ha establecido un plan para garantizar la objetividad de las fuentes de adquisición/recopilación de datos? (Ejemplo de fuentes: personas, instituciones, empresas, etc.) |
| Adquisición/recopilación de datos | * ¿Se han presentado evidencias que respalden la objetividad de las fuentes utilizadas para la adquisición/recopilación de datos? |

<Tabla 46 – Contenido de inspección para la fiabilidad>

* + 1. Suficiencia
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si las categorías (sistema de clasificación) e instancias (datos con características específicas dentro de cada categoría) tienen una cantidad suficiente para ser útiles en el aprendizaje. |
| Elemento sujeto a inspección | Entregables | * Guía de construcción, datos en bruto. |
| Tipo de datos | * Tipos de datos: Texto, imagen, video, audio, 3D. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 47 – Criterios y métodos de inspección para la suficiencia>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de construcción | * ¿Se ha establecido un procedimiento para determinar la cantidad mínima de datos necesarios por sistema de clasificación para el modelo de inteligencia artificial? * ¿Se han presentado evidencias que respalden la determinación de la cantidad mínima de datos por sistema de clasificación para el modelo de inteligencia artificial? |
| Adquisición/recopilación de datos | * ¿Se ha garantizado la cantidad mínima de datos requerida por sistema de clasificación para el modelo de inteligencia artificial? (Ejemplo: el conjunto de datos MS COCO requiere al menos 5,000 instancias por objeto). |

<Tabla 48 – Contenido de inspección para la suficiencia>

* + 1. Uniformidad
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si se ha considerado la uniformidad y proporción en la cantidad de instancias (datos con características específicas) dentro de cada categoría para clasificación, detección, reconocimiento, comprensión y predicción. |
| Elemento sujeto a inspección | Entregables | * Guía de construcción, datos en bruto. |
| Tipo de datos | * Tipos de datos: Texto, imagen, video, audio, 3D. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 49 – Criterios y métodos de inspección para la uniformidad>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de construcción | * ¿Se ha establecido un procedimiento para determinar una proporción adecuada en la cantidad de datos recopilados/adquiridos por sistema de clasificación para el modelo de inteligencia artificial? * ¿Se han presentado evidencias que respalden la determinación de la proporción adecuada en la cantidad de datos recopilados/adquiridos por sistema de clasificación para el modelo de inteligencia artificial? |
| Adquisición/recopilación de datos | * ¿Se ha garantizado que la cantidad de datos recopilados/adquiridos por sistema de clasificación cumple con la proporción adecuada requerida para el modelo de inteligencia artificial? |

<Tabla 50 – Contenido de inspección para la uniformidad>

* + 1. Veracidad
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si la construcción de instancias en cada categoría (clasificación, detección, reconocimiento, comprensión y predicción) refleja entornos y características reales. |
| Elemento sujeto a inspección | Entregables | * Guía de construcción, datos en bruto. |
| Tipo de datos | * Tipos de datos: Texto, imagen, video, audio, 3D. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 51 – Criterios y métodos de inspección para la veracidad>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de construcción | * ¿Se ha establecido un plan para reflejar las características del entorno y las condiciones reales cuando la adquisición/recopilación de datos se realiza en un entorno artificial? * ¿Se ha diseñado un plan para garantizar la consistencia entre las condiciones del entorno real cuando la adquisición/recopilación de datos se lleva a cabo en un entorno artificial? * ¿Se ha establecido un plan para la sincronización entre múltiples fuentes de datos durante la adquisición/recopilación de datos? |
| Adquisición/recopilación de datos | * ¿Se han presentado evidencias que respalden la inclusión de características del entorno y condiciones reales cuando la adquisición/recopilación de datos se ha realizado en un entorno artificial? * ¿Se han proporcionado evidencias que demuestren la consistencia del entorno y condiciones del mundo real en caso de adquisición/recopilación de datos en un entorno artificial? * ¿Se han presentado evidencias de la sincronización entre múltiples fuentes de datos durante la adquisición/recopilación de datos? |

<Tabla 52 – Contenido de inspección para la veracidad>

* + 1. Equidad
       1. Criterios y métodos de inspección

|  |  |  |
| --- | --- | --- |
| Clasificación | | Criterios y métodos de inspección |
| Detalles de la inspección | | * Verificación de si se han eliminado los datos sesgados que podrían estar presentes en el conjunto de datos. |
| Elemento sujeto a inspección | Entregables | * Guía de construcción, datos en bruto. |
| Tipo de datos | * Tipos de datos: Texto, imagen, video, audio, 3D. |
| Método de inspección | | * Método de inspección: Lista de verificación (checklist). |

<Tabla 53 – Criterios y métodos de inspección para la equidad>

* + - 1. Contenido de la inspección

|  |  |
| --- | --- |
| Clasificación | Contenido principal |
| Elaboración del plan de construcción | * ¿Se ha establecido un plan para prevenir sesgos regionales, sociales y raciales en la construcción de datos de entrenamiento de inteligencia artificial? |
| Adquisición/recopilación de datos | * ¿Se han presentado evidencias que respalden la eliminación de sesgos regionales, sociales y raciales en la construcción de datos de entrenamiento de inteligencia artificial? |

<Tabla 54 – Contenido de inspección para la equidad>

**Referencias**

Guía de gestión de calidad de datos de entrenamiento de inteligencia artificial v3.1 (NIA)

Investigación sobre el análisis del estándar de certificación de calidad de datos de inteligencia artificial y la implementación del proceso de certificación de datos (Korea Standards Association)

Marco de ética de datos en inteligencia artificial (Reino Unido)

Guía de gestión de calidad en aprendizaje automático, 3ª edición (Japón)

ISO 8000-2, Calidad de datos (ISO/IEC)

ISO/IEC 25012, Modelo de calidad de datos (ISO/IEC)

Ley de contrataciones del Estado

Ley de Protección de Datos Personales