



# MÁSTER UNIVERSITARIO EN INGENIERÍA DE TELECOMUNICACIÓN

TRABAJO FIN DE MÁSTER

## **Automatización de informes PAI en centros residenciales mediante integración de datos e IA**

Autor: Miguel González Lavín

Director: Alfonso de Lucas David

**Madrid**

Junio de 2026



### Declaración de originalidad

Declaro bajo mi responsabilidad que el Proyecto presentado con el título **Automatización de informes PAI en centros residenciales mediante integración de datos e IA** e la ETS de Ingeniería – ICAI de la Universidad Pontificia Comillas en el curso académico **2025/26** es de mi autoría y no ha sido presentado con anterioridad a otros efectos. El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido tomada de otros documentos está debidamente referenciada.

### Uso de Inteligencia Artificial<sup>1</sup>

Declaro bajo mi responsabilidad que (indicar la opción correcta):

☐ No he utilizado Inteligencia Artificial en la elaboración del presente documento.

☒ He utilizado Inteligencia Artificial en la elaboración del presente documento y/o del Anexo B siempre en las condiciones permitidas por la Universidad Pontificia Comillas, es decir, aplicando el Nivel 2 de la [Escala de Evaluación de Perkins et al. \(2024\)](#): *“La IA puede utilizarse para actividades previas a la tarea, como la lluvia de ideas, la descripción y la investigación inicial. Este nivel se centra en el uso de la IA para la planificación, las síntesis y la generación de ideas, pero las evaluaciones deben hacer hincapié en la capacidad de desarrollar y refinar estas ideas de forma independiente”*. En concreto, las Inteligencia Artificial ha sido empleada para:

La Inteligencia Artificial ha sido empleada como apoyo en las siguientes tareas: (1) generación de ideas y estructuración inicial del proyecto; (2) revisión y mejora estilística de textos redactados previamente por el autor; (3) asistencia en la depuración y revisión de fragmentos de código Python desarrollados por el autor; y (4) síntesis de literatura académica para orientar la búsqueda bibliográfica, siendo las fuentes verificadas y referenciadas de forma independiente por el autor. En todos los casos, el desarrollo conceptual, la toma de decisiones técnicas y la elaboración final del contenido han sido realizados por el autor.



Firmado (alumno): Miguel González Lavín

Fecha: 08/07/2026

---

<sup>1</sup> Esta declaración se refiere al uso de la Inteligencia Artificial generativa para realizar los documentos del Proyecto (Anexo B y Memoria). No aplica a Proyectos donde, por su naturaleza, deban emplear inteligencia artificial como parte de los mismos (aplicación de técnicas de aprendizaje automático, redes neuronales, análisis de datos...)

**Autorización para la entrega del Proyecto**

| El Director del Proyecto    | El co-Director del Proyecto (si aplica) |
|-----------------------------|-----------------------------------------|
|                             |                                         |
| Fdo: Alfonso de Lucas David | Fdo:                                    |
| Fecha: 08/07/2026           | Fecha:                                  |





# MÁSTER UNIVERSITARIO EN INGENIERÍA DE TELECOMUNICACIÓN

TRABAJO FIN DE MÁSTER

## **Automatización de informes PAI en centros residenciales mediante integración de datos e IA**

Autor: Miguel González Lavín

Director: Alfonso de Lucas David

**Madrid**

Junio de 2026

# Agradecimientos

A mis padres, por su apoyo incondicional y por enseñarme desde pequeño que el esfuerzo siempre tiene recompensa.

A mi hermana, por su ánimo constante, por estar presente en cada momento y por ser un referente para mí.

A mi abuela, por todo su cariño y por haberme cuidado siempre con tanta dedicación.

A mis amigos de la universidad, por los momentos vividos y por hacer más llevadera cada etapa del camino.

A todo el equipo de Arpada, en especial al departamento de tecnología, por la ayuda durante todo el proceso.

A todo el equipo de Valdeluz por su acogida y por facilitarme los recursos necesarios para desarrollar este proyecto.

A mi director, Alfonso De Lucas David, por su guía, su implicación y su disposición a lo largo de todo el desarrollo de este trabajo.



# **AUTOMATIZACIÓN DE INFORMES PAI EN CENTROS RESIDENCIALES MEDIANTE INTEGRACIÓN DE DATOS E IA**

**Autor: González Lavín, Miguel.**

Director: De Lucas David, Alfonso.

Entidad Colaboradora: Arpada S. A.

## **RESUMEN DEL PROYECTO**

Este trabajo presenta el diseño e implementación de un sistema asistido para la generación automatizada de informes del Plan de Atención Individualizada (PAI) en centros residenciales. El sistema integra datos estructurados y no estructurados procedentes del ERP ResiPlus y de un datalake corporativo, combinando un pipeline de procesamiento modular con técnicas de Inteligencia Artificial controlada para la síntesis de información clínica.

**Palabras clave:** Inteligencia Artificial, NLP, informes clínicos, entornos residenciales, ResiPlus, datalake

### **1. Introducción**

La transformación digital del sector sociosanitario ha impulsado la adopción de sistemas de información capaces de centralizar y explotar grandes volúmenes de datos clínicos y asistenciales. En los centros residenciales, el seguimiento de las personas mayores se articula principalmente a través del Plan de Atención Individualizada (PAI) [1], documento clave que integra valoraciones médicas, de enfermería, psicológicas y sociales.

A pesar de la implantación de sistemas de gestión como ResiPlus [2], la elaboración de los informes de seguimiento continúa siendo un proceso mayoritariamente manual, basado en la revisión y síntesis de históricos extensos de información. Esta situación genera una elevada carga administrativa y aumenta el riesgo de inconsistencias u omisiones relevantes. En este contexto, la aplicación controlada de técnicas de Inteligencia Artificial orientadas a la síntesis de información clínica se presenta como una oportunidad para optimizar los flujos de trabajo sin sustituir el criterio profesional

### **2. Definición del proyecto**

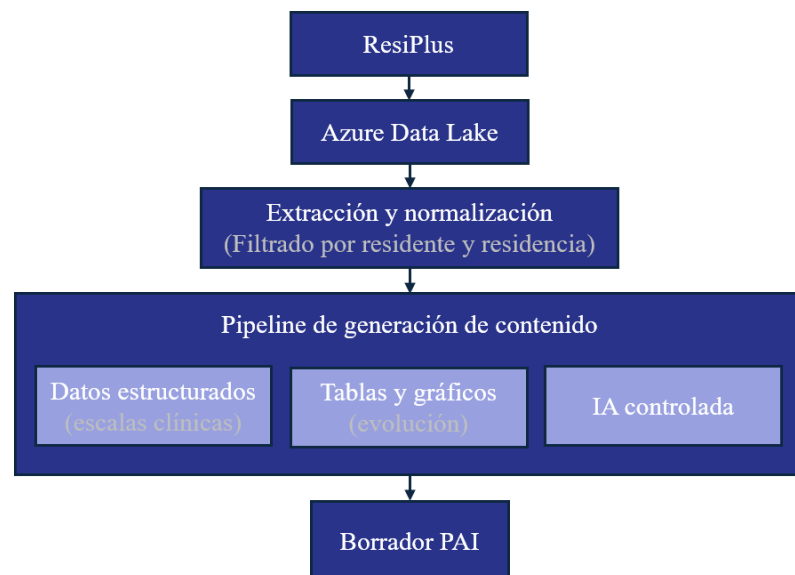
El proyecto tiene como objetivo el diseño y desarrollo de una herramienta asistida para la generación de informes de seguimiento PAI, actuando como un apoyo a los profesionales sanitarios y técnicos de los centros residenciales. La solución se basa en la integración de datos estructurados y no estructurados procedentes del ERP ResiPlus y de un datalake corporativo desplegado sobre tecnologías en la nube [3].

El sistema propuesto no pretende automatizar completamente la toma de decisiones clínicas, sino generar borradores estructurados y coherentes, que posteriormente son revisados y validados por el profesional responsable. De este modo, se garantiza que la Inteligencia Artificial actúe como un “copiloto” de redacción, respetando los requisitos normativos, éticos y de seguridad exigidos en el tratamiento de datos sanitarios

### 3. Descripción del Sistema desarrollado

La arquitectura del sistema se apoya en un pipeline de datos que permite la extracción, transformación y normalización de la información clínica almacenada en el datalake. Estos datos son procesados mediante técnicas de Procesamiento del Lenguaje Natural (NLP) [4] y modelos de lenguaje de gran tamaño, configurados bajo un enfoque de uso controlado, incorporando mecanismos de validación y restricciones para evitar inferencias incorrectas.

La herramienta integra un módulo de generación documental que permite volcar automáticamente la información sintetizada en plantillas Word predefinidas, adaptadas a la estructura del PAI. La conexión entre los distintos módulos garantiza un flujo de trabajo continuo, desde la obtención de datos hasta la generación del borrador final del informe.



*Ilustración 1 - Sistema desarrollado*

### 4. Resultados

La validación del sistema se ha realizado mediante pruebas funcionales y escenarios reales de uso, comparando el proceso tradicional de elaboración manual de informes con el proceso asistido propuesto. Los resultados obtenidos evidencian una reducción significativa del tiempo dedicado a la redacción de informes, así como una mejora en la coherencia estructural y terminológica de los documentos generados.

Asimismo, el feedback de los profesionales participantes destaca la utilidad del sistema como herramienta de apoyo, permitiendo centrar los esfuerzos en la validación clínica y la toma de decisiones, en lugar de en tareas administrativas repetitivas.

El sistema ha sido validado sobre casos reales de residentes de los centros Valdeluz, confirmando su viabilidad técnica y su adecuación a los requisitos normativos y organizativos del entorno sociosanitario.

## 5. Conclusiones

El trabajo desarrollado demuestra que la aplicación controlada de técnicas de Inteligencia Artificial, integrada con sistemas de gestión y arquitecturas de datos en la nube, puede aportar un valor real y medible en el ámbito sociosanitario. La solución propuesta mejora la eficiencia operativa sin comprometer la seguridad, la calidad asistencial ni el cumplimiento normativo.

Como principal aportación, el proyecto establece un enfoque viable para la automatización asistida de informes clínicos, alineado con las necesidades reales de los centros residenciales. Este trabajo abre la puerta a futuras ampliaciones del sistema, como la incorporación de nuevos indicadores clínicos, mejoras en la personalización de los informes o la extensión a otros ámbitos asistenciales.

## 6. Referencias

- [1] Boletín Oficial de la Comunidad de Madrid (n.º 273, 16 de noviembre 2024). Orden por la que se regula el Plan de Atención Individualizada (PAI) en centros residenciales [https://www.bocm.es/boletin/CM\\_Orden\\_BOCM/2024/11/16/BOCM-20241116-1.PDF](https://www.bocm.es/boletin/CM_Orden_BOCM/2024/11/16/BOCM-20241116-1.PDF).
- [2] Resiplus. Software de gestión para residencias y centros sociosanitarios <https://addinformatica.com/>.
- [3] Microsoft Azure. Azure Data Lake: Solución de almacenamiento de datos a gran escala. <https://azure.microsoft.com/es-es/solutions/data-lake>.
- [4] The Bridge. NLP: Procesamiento del Lenguaje Natural y sus aplicaciones. <https://thebridge.tech/blog/nlp-procesamiento-lenguaje-natural/>.

# **AUTOMATION OF PAI REPORTS IN RESIDENTIAL CENTERS THROUGH DATA INTEGRATION AND AI**

**Author: González Lavín, Miguel.**

Supervisor: De Lucas David, Alfonso.

Collaborating Entity: Arpada S. A.

## **ABSTRACT**

This work presents the design and implementation of an assisted system for the automated generation of Individual Care Plan (PAI) reports in residential care centres. The system integrates structured and unstructured data from the ResiPlus ERP and a corporate data lake, combining a modular processing pipeline with controlled Artificial Intelligence techniques for the synthesis of clinical information.

**Keywords:** Artificial Intelligence, NLP, clinical reports, residential care settings, ResiPlus, data lake

## **1. Introduction**

The digital transformation of the socio-healthcare sector has driven the adoption of information systems capable of centralising and exploiting large volumes of clinical and care-related data. In residential centres, the monitoring of elderly residents is primarily structured around the Individual Care Plan (PAI) [1], a key document integrating medical, nursing, psychological and social assessments.

Despite the deployment of management systems such as ResiPlus [2], the preparation of follow-up reports remains a largely manual process, based on the review and synthesis of extensive historical records. This situation generates a significant administrative burden and increases the risk of inconsistencies or relevant omissions. In this context, the controlled application of Artificial Intelligence techniques oriented towards the synthesis of clinical information represents an opportunity to optimise workflows without replacing professional judgement.

## **2. Project definition**

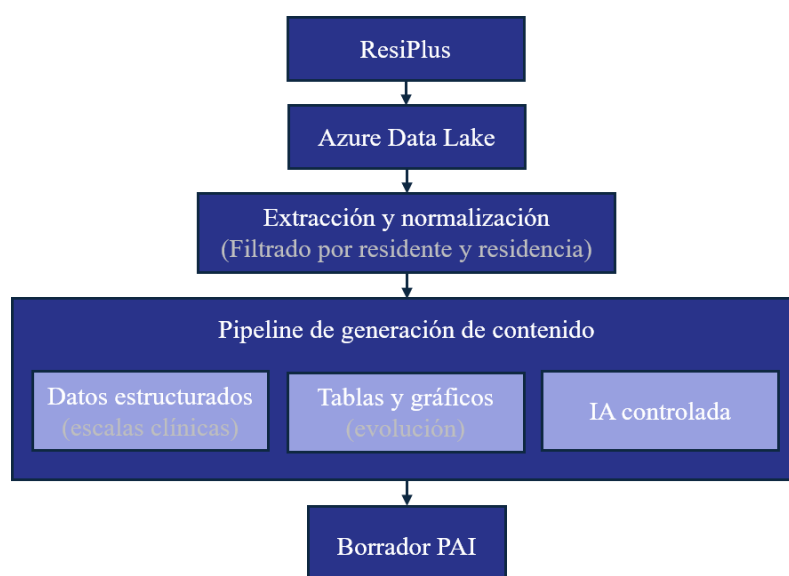
The project aims to design and develop an assisted tool for the generation of PAI follow-up reports, acting as a support for healthcare and technical professionals in residential centres. The solution is based on the integration of structured and unstructured data from the ResiPlus ERP and a corporate data lake deployed on cloud technologies [3].

The proposed system does not seek to fully automate clinical decision-making, but rather to generate structured and coherent drafts, which are subsequently reviewed and validated by the responsible professional. In this way, Artificial Intelligence is ensured to act as a drafting "co-pilot", complying with the regulatory, ethical and security requirements applicable to the processing of healthcare data.

### 3. Description of the developed system

The system architecture is built upon a data pipeline that enables the extraction, transformation and normalisation of clinical information stored in the data lake. This data is processed using Natural Language Processing (NLP) techniques [4] and large language models, configured under a controlled-use approach, incorporating validation mechanisms and constraints to prevent incorrect inferences.

The tool integrates a document generation module that automatically transfers the synthesised information into predefined Word templates, adapted to the PAI structure. The connection between the different modules ensures a continuous workflow, from data retrieval through to the generation of the final report draft.



*Ilustración 2 - System structure*

### 4. Results

The system has been validated through functional tests and real-use scenarios, comparing the traditional manual report preparation process with the proposed assisted process. The results obtained demonstrate a significant reduction in the time devoted to report drafting, as well as an improvement in the structural and terminological consistency of the generated documents.

Furthermore, feedback from the participating professionals highlights the usefulness of the system as a support tool, allowing efforts to be focused on clinical validation and decision-making rather than on repetitive administrative tasks.

The system has been validated on real resident cases from the Valdeluz centres, confirming its technical feasibility and its suitability for the regulatory and organisational requirements of the socio-healthcare environment.

## 5. Conclusions

The work carried out demonstrates that the controlled application of Artificial Intelligence techniques, integrated with management systems and cloud-based data architectures, can deliver real and measurable value in the socio-healthcare sector. The proposed solution improves operational efficiency without compromising safety, care quality or regulatory compliance.

As its main contribution, the project establishes a viable approach for the assisted automation of clinical reports, aligned with the real needs of residential care centres. This work opens the door to future extensions of the system, such as the incorporation of new clinical indicators, improvements in report personalisation, or expansion to other care settings.

## 6. References

- [1] Boletín Oficial de la Comunidad de Madrid (No. 273, 16 November 2024). Order regulating the Individual Care Plan (PAI) in residential centres.  
[https://www.bocm.es/boletin/CM\\_Orden\\_BOCM/2024/11/16/BOCM-20241116-1.PDF](https://www.bocm.es/boletin/CM_Orden_BOCM/2024/11/16/BOCM-20241116-1.PDF).
- [2] Resiplus. Management software for residential and socio-healthcare centres.  
<https://addinformatica.com/>.
- [3] Microsoft Azure. Azure Data Lake: Large-scale data storage solution.  
<https://azure.microsoft.com/es-es/solutions/data-lake>.
- [4] The Bridge. NLP: Natural Language Processing and its applications.  
<https://thebridge.tech/blog/nlp-procesamiento-lenguaje-natural/>.

## *Índice de la memoria*

|                                                                                |           |
|--------------------------------------------------------------------------------|-----------|
| <b>Capítulo 1. Introducción .....</b>                                          | <b>8</b>  |
| <b>Capítulo 2. Descripción de las tecnologías.....</b>                         | <b>10</b> |
| 2.1 Lenguajes de programación.....                                             | 10        |
| 2.2 Entorno de desarrollo y control de versiones .....                         | 11        |
| 2.3 Fuentes de datos: ResiPlus .....                                           | 11        |
| 2.4 Almacenamiento y arquitectura de datos .....                               | 12        |
| 2.5 Tecnologías de Inteligencia Artificial .....                               | 12        |
| <b>Capítulo 3. Estado de la Cuestión.....</b>                                  | <b>14</b> |
| 3.1 Evolución de la atención sociosanitaria .....                              | 14        |
| 3.2 Contexto del Plan de Atención Individualizada (PAI) .....                  | 17        |
| 3.3 Sistemas de gestión sociosanitaria actuales.....                           | 18        |
| 3.4 Automatización y generación de informes en el ámbito clínico .....         | 20        |
| 3.5 Uso de Inteligencia Artificial en la síntesis de información clínica ..... | 21        |
| <b>Capítulo 4. Definición del Trabajo .....</b>                                | <b>23</b> |
| 4.1 Justificación.....                                                         | 23        |
| 4.2 Objetivos .....                                                            | 24        |
| 4.3 Metodología.....                                                           | 26        |
| 4.4 Planificación y Estimación Económica .....                                 | 27        |
| <b>Capítulo 5. Sistema Desarrollado .....</b>                                  | <b>34</b> |
| <b>Capítulo 6. Datos de entrada y tratamiento de la información.....</b>       | <b>36</b> |
| 6.1 Origen de los datos .....                                                  | 36        |
| 6.2 Tipología de los datos tratados .....                                      | 37        |
| 6.2.1 Datos estructurados: escalas clínicas y registros cuantitativos.....     | 37        |
| 6.2.2 Datos no estructurados: notas y observaciones clínicas.....              | 38        |
| 6.2.3 Datos temporales: registros evolutivos .....                             | 39        |
| 6.3 Procesos de extracción de datos .....                                      | 40        |
| 6.4 Tratamiento, normalización y anonimización de la información .....         | 41        |

|                                                                      |           |
|----------------------------------------------------------------------|-----------|
| <b>Capítulo 7. Arquitectura y estructura del programa.....</b>       | <b>42</b> |
| 7.1 Visión general de la arquitectura.....                           | 42        |
| 7.2 Flujo de ejecución del sistema.....                              | 43        |
| 7.3 Estructura del proyecto.....                                     | 43        |
| 7.4 Módulo pipeline .....                                            | 45        |
| 7.5 Módulo sections .....                                            | 45        |
| 7.6 Módulos auxiliares .....                                         | 46        |
| 7.7 Recursos y configuración externa .....                           | 46        |
| 7.8 Escalabilidad, mantenimiento y control del sistema .....         | 47        |
| <b>Capítulo 8. Generación de informes y plantillas.....</b>          | <b>48</b> |
| 8.1 Diseño del modelo de generación documental .....                 | 48        |
| 8.2 Justificación del uso de plantillas estructuradas .....          | 49        |
| 8.3 Sistema de etiquetas dinámicas .....                             | 50        |
| 8.4 Sustitución de contenido y gestión de excepciones.....           | 51        |
| 8.4.1 Sustitución por texto simple.....                              | 51        |
| 8.4.2 Sustitución por tablas dinámicas.....                          | 53        |
| 8.4.3 Sustitución por gráficas.....                                  | 56        |
| 8.4.4 Sustitución por bloques generados mediante IA .....            | 57        |
| 8.4.5 Gestión de excepciones documentales.....                       | 58        |
| <b>Capítulo 9. Modelo controlado de Inteligencia Artificial.....</b> | <b>60</b> |
| 9.1 Principios del uso controlado de la IA .....                     | 60        |
| 9.2 Diseño y estructuración de prompts .....                         | 61        |
| 9.3 Ejemplos de prompts utilizados.....                              | 61        |
| <b>Capítulo 10. Análisis de Resultados.....</b>                      | <b>64</b> |
| 10.1 Validación funcional del sistema.....                           | 64        |
| 10.2 Evaluación de eficiencia y tiempo.....                          | 65        |
| 10.3 Valoración por parte de los profesionales .....                 | 66        |
| 10.4 Limitaciones detectadas .....                                   | 67        |
| <b>Capítulo 11. Conclusiones y Trabajos Futuros.....</b>             | <b>68</b> |
| 11.1 Conclusiones .....                                              | 68        |
| 11.2 Trabajos futuros.....                                           | 69        |

---

|                                                                   |           |
|-------------------------------------------------------------------|-----------|
| <b>Capítulo 12. Bibliografía.....</b>                             | <b>71</b> |
| <b>ANEXO I: Alineación del proyecto con los ODS.....</b>          | <b>75</b> |
| <b>ANEXO II: Librerías utilizadas.....</b>                        | <b>79</b> |
| <b>ANEXO III: Escalas clínicas.....</b>                           | <b>81</b> |
| III.1 Escala de Lobo (MEC - Mini-Examen Cognoscitivo).....        | 81        |
| III.2 Índice de Barthel.....                                      | 84        |
| III.3 Escala de Tinetti.....                                      | 86        |
| III.4 Escala de Norton.....                                       | 88        |
| III.5 Escala de Downton.....                                      | 89        |
| III.6 Timed Up & Go (TUG).....                                    | 89        |
| III.7 Escala GDS (Global Deterioration Scale).....                | 90        |
| III.8 Escala EVA (Escala Visual Analógica).....                   | 91        |
| III.9 Escala Yesavage (GDS-15).....                               | 91        |
| III.10 Inventario Neuropsiquiátrico (NPI).....                    | 93        |
| III.11 Escala Mini Nutritional Assessment (MNA).....              | 94        |
| III.12 Escala PAINAD (Paint Assessment in Advanced Dementia)..... | 97        |

## *Índice de figuras*

|                                                                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1 - Pirámide de población para el año 2022. Elaboración propia a partir de información publicada por el Instituto Nacional de Estadística (INE)..... | 15 |
| Figura 2 - Pirámide de población para el año 2052. Elaboración propia a partir de información publicada por el Instituto Nacional de Estadística (INE)..... | 15 |
| Figura 3 - Pirámide de población para el año 2072. Elaboración propia a partir de información publicada por el Instituto Nacional de Estadística (INE)..... | 16 |
| Figura 4 - Diagrama de Gantt del proyecto .....                                                                                                             | 28 |
| Figura 5 - Arquitectura del sistema .....                                                                                                                   | 34 |
| Figura 6 - Proceso de extracción y tratamiento de la información .....                                                                                      | 37 |
| Figura 7 - Ejemplo datos no estructurados .....                                                                                                             | 39 |
| Figura 8 - Ejemplo evolutivo Norton .....                                                                                                                   | 39 |
| Figura 9 - Ejemplo de sustitución por texto simple.....                                                                                                     | 52 |
| Figura 10 - Ejemplo de sustitución por tabla dinámica .....                                                                                                 | 55 |
| Figura 11 - Ejemplo de sustitución por gráfica .....                                                                                                        | 57 |
| Figura 12 / Ejemplo de texto generado por IA .....                                                                                                          | 58 |
| Figura 13 - Escala EVA.....                                                                                                                                 | 91 |

## *Índice de tablas*

|                                                                               |    |
|-------------------------------------------------------------------------------|----|
| Tabla 1 - Costes asociados al equipo que desarrolla el proyecto .....         | 29 |
| Tabla 2 - Costes asociados a los servicios utilizados .....                   | 30 |
| Tabla 3 - Costes por uso de los modelos de Inteligencia Artificial .....      | 30 |
| Tabla 4 - Costes de mantenimiento y explotación Año 0 + 5 primeros años ..... | 31 |
| Tabla 5 - Ejemplo datos estructurados .....                                   | 38 |
| Tabla 6 - Estructura del proyecto .....                                       | 44 |
| Tabla 7 - Listado de los ODS .....                                            | 75 |
| Tabla 8 - Librerías utilizadas .....                                          | 79 |
| Tabla 9 - Preguntas Escala Lobo de Orientación temporal .....                 | 82 |
| Tabla 10 - Preguntas Escala Lobo de Orientación espacial.....                 | 82 |
| Tabla 11 - Preguntas Escala Lobo de Fijación .....                            | 82 |
| Tabla 12 - Preguntas Escala Lobo de Concentración y cálculo .....             | 83 |
| Tabla 13 - Preguntas Escala Lobo de Memoria.....                              | 83 |
| Tabla 14 - Preguntas Escala Lobo de Lenguaje y construcción .....             | 83 |
| Tabla 15 - Preguntas Índice de Barthel .....                                  | 85 |
| Tabla 16 - Preguntas Escala de Tinetti para la evaluación del equilibrio..... | 86 |
| Tabla 17 - Preguntas Escala de Tinetti para la evaluación de la marcha.....   | 87 |
| Tabla 18 - Preguntas escala de Norton .....                                   | 88 |
| Tabla 19 - Descripción estados en la escala GDS .....                         | 90 |
| Tabla 20 - Preguntas escala Yesavage.....                                     | 92 |
| Tabla 21 - Síntomas evaluados en NPI.....                                     | 93 |
| Tabla 22 - Preguntas escala de Mini Nutritional Assessment I.....             | 95 |
| Tabla 23 - Preguntas escala de Mini Nutritional Assessment II .....           | 96 |

## *Índice de abreviaturas*

| <i>Abreviatura</i> | <i>Significado</i>                                                                |
|--------------------|-----------------------------------------------------------------------------------|
| ABVD               | Actividades Básicas de la Vida Diaria                                             |
| ADLS               | Azure Data Lake Storage                                                           |
| API                | Application Programming Interface (Interfaz de Programación de Aplicaciones)      |
| BOE                | Boletín Oficial del Estado                                                        |
| CSV                | Comma-Separated Values (Formato de datos separados por comas)                     |
| ERP                | Enterprise Resource Planning (Sistema de Planificación de Recursos Empresariales) |
| EVA                | Escala Visual Analógica                                                           |
| GDS                | Global Deterioration Scale                                                        |
| IA                 | Inteligencia Artificial                                                           |
| INE                | Instituto Nacional de Estadística                                                 |
| LLM                | Large Language Model (Modelo de Lenguaje de Gran Tamaño)                          |
| MEC                | Mini-Examen Cognoscitivo                                                          |
| MNA                | Mini Nutritional Assessment                                                       |

|        |                                                                  |
|--------|------------------------------------------------------------------|
| NLP    | Natural Language Processing (Procesamiento del Lenguaje Natural) |
| NPI    | Inventario Neuropsiquiátrico                                     |
| OCDE   | Organización para la Cooperación y el Desarrollo Económicos      |
| OMS    | Organización Mundial de la Salud                                 |
| ODS    | Objetivos de Desarrollo Sostenible                               |
| PAI    | Plan de Atención Individualizada                                 |
| PAINAD | Pain Assessment in Advanced Dementia                             |
| SQL    | Structured Query Language                                        |
| TUG    | Timed Up & Go                                                    |
| UE     | Unión Europea                                                    |

## Capítulo 1. INTRODUCCIÓN

En la actualidad, el sector sociosanitario se encuentra inmerso en un proceso de transformación impulsado por el envejecimiento progresivo de la población, el aumento de la prevalencia de enfermedades crónicas y la creciente necesidad de ofrecer una atención integral y personalizada a las personas mayores. Este contexto plantea importantes retos organizativos y asistenciales, especialmente en los centros residenciales, donde la calidad del cuidado depende en gran medida de la correcta gestión y explotación de la información clínica y asistencial.

Uno de los elementos clave es el Plan de Atención Individualizada (PAI), documento que recoge de forma estructurada la situación clínica, funcional, cognitiva, social y emocional de cada residente, así como las intervenciones realizadas por los distintos profesionales implicados en su cuidado. Este constituye una herramienta fundamental para garantizar una atención centrada en la persona y un seguimiento adecuado de su evolución. Sin embargo, a pesar de la implantación generalizada de sistemas de gestión sociosanitaria que permiten registrar grandes volúmenes de información, la elaboración de estos informes continúa siendo, en muchos casos, un proceso manual y laborioso. Los profesionales deben revisar extensos historiales de notas clínicas, escalas de valoración y registros asistenciales dispersos para sintetizar la información relevante en documentos estructurados, lo que supone una carga administrativa significativa. Esto desemboca en dos problemas principales: la inversión de tiempo necesaria y la posibilidad de errores humanos.

En primer lugar, el elevado consumo de tiempo dedicado a tareas de síntesis documental se detrae de la atención directa al residente y del análisis clínico, actividades que aportan un mayor valor asistencial. En segundo lugar, la elaboración manual de informes incrementa el riesgo de omisiones, inconsistencias y falta de homogeneidad en los documentos generados, especialmente cuando intervienen múltiples perfiles profesionales y se manejan grandes volúmenes de información histórica.

En este contexto, el presente Trabajo de Fin de Máster “Automatización de informes PAI en centros residenciales mediante integración de datos e IA” propone abordar estas limitaciones mediante el desarrollo de un sistema de generación asistida y automatizada de informes de seguimiento del Plan de Atención Individualizada. La solución planteada se basa en la integración avanzada de los datos procedentes del sistema de gestión residencial ResiPlus y del datalake corporativo, junto con el uso controlado de técnicas de Inteligencia Artificial aplicadas a la síntesis de información clínica. El objetivo del sistema no es sustituir el criterio del profesional sociosanitario, sino proporcionarle una herramienta de apoyo que genere borradores estructurados de los informes PAI a partir de la información ya registrada. De este modo, el profesional puede centrar su labor en la validación clínica, la toma de decisiones y la atención directa al residente, reduciendo la carga administrativa asociada a tareas repetitivas de redacción y síntesis.

El trabajo que aquí se expone se estructura de la siguiente manera. En el Capítulo 2 se describen las tecnologías empleadas en el desarrollo del proyecto, incluyendo los lenguajes de programación, las arquitecturas de datos y las técnicas de Inteligencia Artificial utilizadas. En el Capítulo 3 se analiza el estado de la cuestión, abordando la evolución de la atención sociosanitaria, el contexto del Plan de Atención Individualizada y las soluciones existentes en el ámbito de la automatización de informes clínicos. Posteriormente, en el Capítulo 4 se define el trabajo desarrollado, junto con la justificación, objetivos, metodología y planificación económica.

A continuación, los Capítulos 5 a 9 detallan el sistema desarrollado, describiendo las fuentes de datos, la arquitectura del programa, el proceso de generación de informes y el modelo controlado de Inteligencia Artificial implementado. En el Capítulo 10 se analizan los resultados obtenidos y se evalúa el impacto funcional y organizativo del sistema. Finalmente, en el Capítulo 11 se recogen las conclusiones del proyecto y se plantean las líneas de trabajo futuro.

## **Capítulo 2. DESCRIPCIÓN DE LAS TECNOLOGÍAS**

En este capítulo se describen las principales tecnologías, herramientas y enfoques empleados en el desarrollo de este Trabajo de Fin de Máster. El objetivo es facilitar la lectura y comprensión de los capítulos posteriores, proporcionando el contexto técnico necesario. La arquitectura del proyecto combina herramientas de extracción y almacenamiento de datos en la nube, procesamiento y modelado en Python, y el control de versiones mediante Bitbucket.

### **2.1 LENGUAJES DE PROGRAMACIÓN**

El desarrollo del sistema se ha realizado utilizando Python [1] como lenguaje de programación principal. Su elección se fundamenta en su versatilidad, claridad sintáctica y amplio uso en proyectos relacionados con el tratamiento de datos, la automatización de procesos y la integración con tecnologías de Inteligencia Artificial.

Python dispone de un ecosistema muy amplio de librerías especializadas que facilitan la manipulación de datos, la conexión con servicios externos y la generación automatizada de documentos. Estas características permiten desarrollar soluciones complejas de forma eficiente y mantenible, adaptándose a los requisitos técnicos del proyecto. El listado completo de librerías utilizadas, junto con una breve descripción de su función, se recoge en el ANEXO II: Librerías utilizadas.

Por otro lado, se ha utilizado el lenguaje SQL [2] para la consulta y explotación de los datos almacenados en las distintas tablas y estructuras del sistema. Permite realizar operaciones de selección, filtrado, ordenación y agregación de datos de manera eficiente.

En el proyecto ha sido fundamental para la extracción de los datos relevantes desde el datalake y otros sistemas de almacenamiento, proporcionando los conjuntos de datos necesarios para su posterior transformación y análisis.

## **2.2 ENTORNO DE DESARROLLO Y CONTROL DE VERSIONES**

Visual Studio Code [3] es el entorno de desarrollo integrado utilizado a lo largo de todo el proyecto. Su elección se justifica por la combinación de ligereza, extensibilidad y soporte nativo para Python y Jupyter Notebooks, que permite trabajar tanto con scripts modulares como con cuadernos de experimentación dentro del mismo entorno. Las extensiones de Python, Pylance [4] y Jupyter [5] proporcionan autocompletado, análisis estático de tipos y ejecución interactiva de celdas, lo que agiliza el ciclo de desarrollo y depuración del código.

El control de versiones del proyecto se gestiona mediante Bitbucket [6], plataforma de alojamiento de repositorios Git. El uso de Git [7] como sistema de control de versiones garantiza la trazabilidad completa de todos los cambios realizados en el código a lo largo del desarrollo, permitiendo revertir modificaciones, comparar versiones y mantener ramas separadas para la experimentación y la versión estable del sistema. Bitbucket se integra de forma nativa con el ecosistema Atlassian, lo que facilita su adopción en entornos corporativos donde otras herramientas del mismo ecosistema ya están implantadas.

La estructura del repositorio sigue una organización modular donde cada fase del pipeline tiene su propio módulo: extracción, preprocesado y reemplazo de datos. Esta separación facilita la reutilización de componentes y la incorporación futura del repositorio a un pipeline de integración continua en el entorno productivo Databricks [8].

## **2.3 FUENTES DE DATOS: RESIPLUS**

ResiPlus [9] es un sistema de planificación de recursos empresariales (ERP) ampliamente implantado en el ámbito sociosanitario para la gestión integral de centros residenciales. Centraliza la información relacionada con los residentes, incluyendo datos administrativos, clínicos, funcionales y asistenciales, así como las anotaciones realizadas por los distintos profesionales implicados en su atención.

En el contexto del proyecto, ResiPlus actúa como sistema origen de los datos utilizados para la elaboración de los Planes de Atención Individualizada (PAI). En él se registra la

información generada en la práctica asistencial diaria, constituyendo la base sobre la que se realiza el análisis y la posterior síntesis asistida de los informes de seguimiento. Esto permite disponer de un histórico estructurado de la evolución del residente.

## ***2.4 ALMACENAMIENTO Y ARQUITECTURA DE DATOS***

La fuente primaria de datos del proyecto es el datalake corporativo, en el que se vuelcan los datos procedentes de ResiPlus, el sistema de gestión sociosanitaria utilizado en los centros residenciales. Este repositorio centraliza el histórico clínico y asistencial de los residentes, incluyendo notas de evolución, escalas de valoración y registros de las distintas disciplinas profesionales implicadas en su atención..

La conexión desde Python se establece mediante el protocolo JDBC [10] usando el driver oficial de Microsoft para SQL Server, gestionado a través de las librerías JPytype [11] y JayDeBeAPI [12]. Esta aproximación permite ejecutar consultas SQL arbitrarias desde el entorno Python sin necesidad de exportaciones manuales intermedias, garantizando que los datos utilizados en cada ejecución del pipeline corresponden siempre a la versión más actualizada disponible en el sistema. Las credenciales de acceso se almacenan en un fichero de configuración externo al repositorio para evitar su exposición en el control de versiones.

## ***2.5 TECNOLOGÍAS DE INTELIGENCIA ARTIFICIAL***

La generación asistida de informes se apoya en tecnologías de Inteligencia Artificial orientadas al tratamiento y síntesis de información textual. Dado el carácter sensible del entorno sociosanitario, estas tecnologías se utilizan bajo un enfoque controlado, priorizando la fiabilidad, la trazabilidad y la validación humana de los resultados generados.

El componente principal es el uso de Modelos de Lenguaje de Gran Tamaño (LLM) [13], sistemas entrenados sobre grandes volúmenes de texto capaces de generar contenido lingüístico coherente y contextualizado. En este proyecto, los LLM se emplean para generar borradores interpretativos a partir de información clínica previamente procesada, actuando

como herramienta de apoyo para los profesionales y manteniendo siempre la supervisión humana sobre el resultado final.

Para garantizar que el contenido generado se base exclusivamente en datos reales, el sistema restringe de forma estricta el contexto de actuación del modelo. La información que recibe el LLM proviene directamente del datalake corporativo, estructurada mediante consultas SQL y procesada en Python antes de ser incorporada al prompt. Este enfoque de recuperación estructurada de información reduce el riesgo de generación de contenido no fundamentado, limitando la actuación del modelo a la interpretación y síntesis de los datos disponibles.

El uso de la Inteligencia Artificial se ha planteado desde un enfoque asistido y controlado, en el que la validación final de los informes corresponde siempre al profesional responsable. La IA actúa como herramienta de apoyo a la redacción, sin sustituir el criterio clínico ni la toma de decisiones, asegurando una integración responsable en los procesos asistenciales existentes.

## Capítulo 3. ESTADO DE LA CUESTIÓN

En este capítulo se analiza el estado actual del ámbito sociosanitario, examinando la evolución de los modelos de atención en centros residenciales y las soluciones tecnológicas aplicadas en este sector. Este análisis permite identificar las limitaciones existentes en los procesos actuales de gestión y seguimiento asistencial, así como las oportunidades de mejora sobre las que se fundamenta el presente proyecto.

Asimismo, se incorpora una breve referencia al contexto internacional, con el fin de situar la evolución del modelo español dentro de una tendencia global hacia sistemas de atención integrados y centrados en la persona.

### 3.1 EVOLUCIÓN DE LA ATENCIÓN SOCIO SANITARIA

Desde finales de la década de 1980, el concepto de atención sociosanitaria se ha consolidado progresivamente en España y en el conjunto de la Unión Europea, impulsado por el envejecimiento sostenido de la población y el fortalecimiento de los sistemas de protección social. La creciente prevalencia de enfermedades crónicas, situaciones de dependencia y discapacidad ha evidenciado la necesidad de ofrecer cuidados integrales que combinen intervenciones sanitarias y sociales de forma coordinada [14].

Las previsiones demográficas han reforzado esta necesidad. Según estimaciones del Instituto Nacional de Estadística (INE): “El porcentaje de población de 65 años y más, que actualmente se sitúa en el 20,4% del total, alcanzaría un máximo del 30,5% en torno a 2055. En 2074 este porcentaje se situaría en el 30,3%” (INE 2024) [15] (Figura 1, Figura 2 y Figura 3). Este crecimiento se acompaña de un aumento significativo del grupo de personas de edad muy avanzada, lo que incrementa la demanda de cuidados de larga duración y de atención especializada. En este contexto, los modelos fragmentados de atención resultan insuficientes para dar respuesta a necesidades complejas que abarcan dimensiones clínicas, funcionales y sociales de forma simultánea.

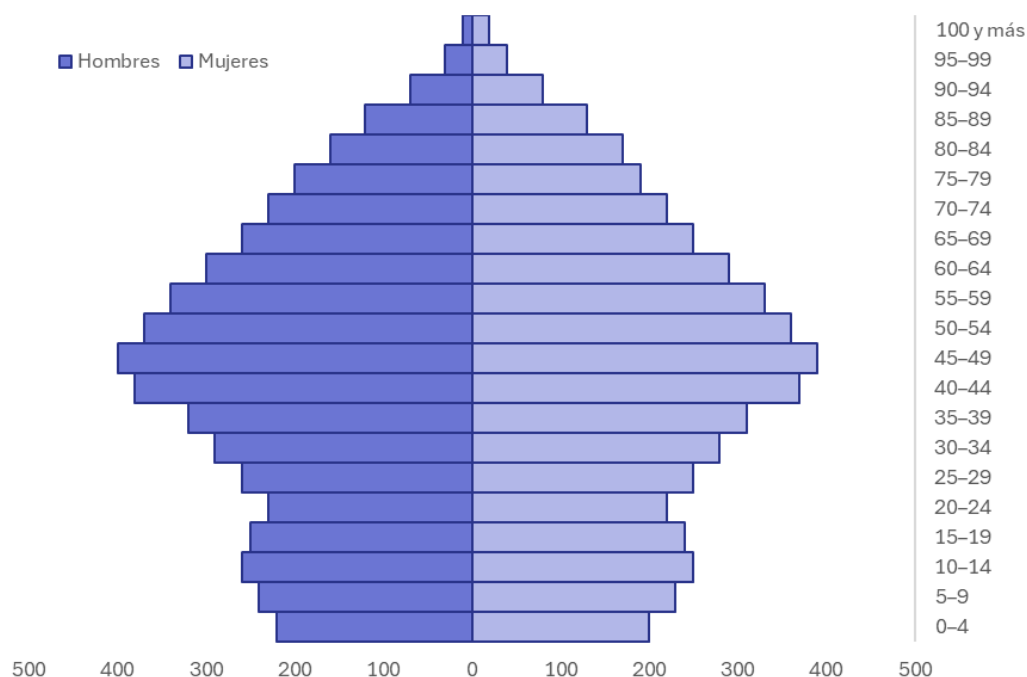


Figura 1 - Pirámide de población para el año 2022. Elaboración propia a partir de información publicada por el Instituto Nacional de Estadística (INE).

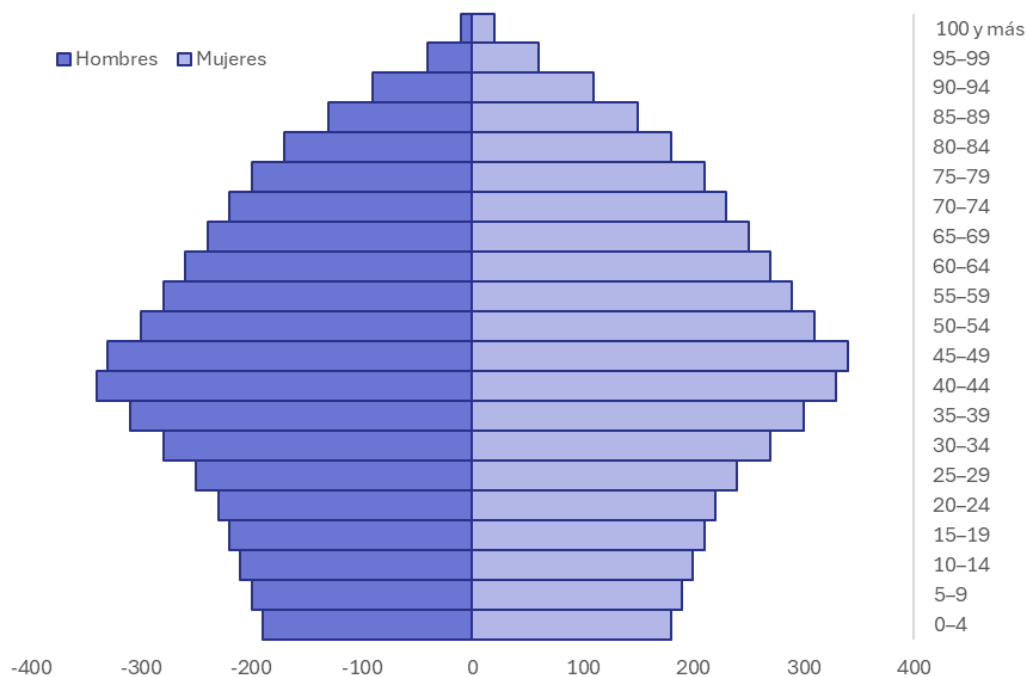
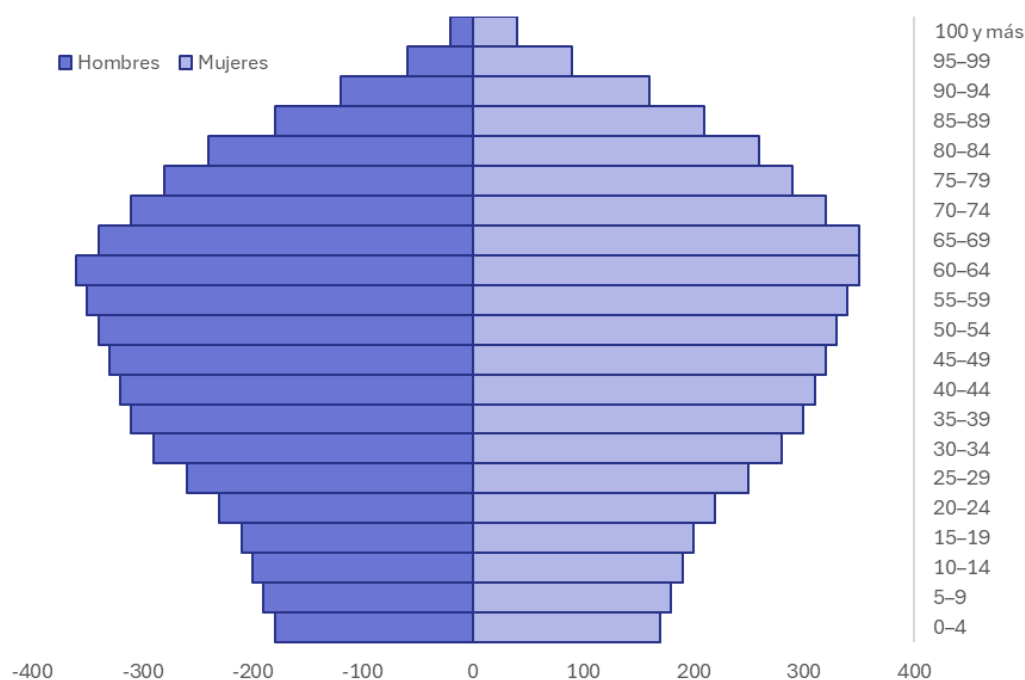


Figura 2 - Pirámide de población para el año 2052. Elaboración propia a partir de información publicada por el Instituto Nacional de Estadística (INE).



*Figura 3 - Pirámide de población para el año 2072. Elaboración propia a partir de información publicada por el Instituto Nacional de Estadística (INE).*

A nivel internacional, la evolución hacia modelos integrados de atención ha seguido trayectorias diversas. En países nórdicos como Suecia o Dinamarca, tradicionalmente se ha desarrollado un modelo fuerte de servicios sociales municipales, con una clara orientación hacia los cuidados de larga duración en el entorno comunitario. En el Reino Unido, el sistema de National Health Service (NHS) ha impulsado progresivamente estrategias de integración entre servicios sanitarios y sociales, especialmente tras la introducción de programas de integrated care systems orientados a mejorar la coordinación territorial [16]. En Alemania, la creación en 1995 del seguro obligatorio de dependencia (Pflegeversicherung) supuso un hito en la estructuración de un sistema específico de financiación de cuidados de larga duración, diferenciando claramente las prestaciones sanitarias de las sociales, pero estableciendo mecanismos de coordinación entre ambas [17].

En términos generales, los informes de la OCDE y de la Organización Mundial de la Salud coinciden en señalar que los sistemas más eficientes son aquellos que avanzan hacia modelos de atención centrados en la persona [18], con mecanismos formales de coordinación

interdisciplinar y herramientas estructuradas de planificación individualizada. La integración asistencial no implica necesariamente la unificación institucional, sino la articulación efectiva de procesos, información y responsabilidades entre los distintos niveles de atención.

En España, aunque la atención sociosanitaria ha sido reconocida como un objetivo estratégico desde principios de los años noventa, su desarrollo ha estado condicionado por la fragmentación competencial entre administraciones autonómicas y por diferencias organizativas y culturales entre los ámbitos sanitario y social. La aprobación de la Ley 39/2006 [19] marcó un punto de inflexión al reconocer de forma explícita el derecho subjetivo a recibir apoyos adaptados a la situación individual de cada persona.

Este contexto normativo y demográfico ha favorecido la consolidación progresiva de herramientas orientadas a la planificación individualizada de la atención y al seguimiento continuado de los usuarios en centros residenciales. Entre estas herramientas destaca el Plan de Atención Individualizada, cuyo marco y función se desarrollan en el apartado siguiente.

### ***3.2 CONTEXTO DEL PLAN DE ATENCIÓN INDIVIDUALIZADA (PAI)***

El Plan de Atención Individualizada (PAI) constituye el instrumento operativo central de la atención sociosanitaria en los centros residenciales. Se configura como un documento dinámico que integra de forma estructurada la información clínica, funcional, cognitiva, social y emocional de cada residente, así como los objetivos asistenciales y las intervenciones planificadas por el equipo interdisciplinar. Su finalidad es garantizar una atención personalizada, coherente y evaluable en el tiempo.

El desarrollo del PAI se enmarca en la evolución normativa y organizativa orientada a consolidar modelos de atención centrados en la persona. En España, este enfoque encuentra un respaldo normativo relevante en la Ley 39/2006 [19], de Promoción de la Autonomía Personal y Atención a las Personas en Situación de Dependencia, que establece la necesidad de diseñar intervenciones ajustadas a las necesidades individuales, promoviendo la

autonomía, la dignidad y la calidad de vida de las personas en situación de dependencia. Aunque la Ley no regula de manera específica el documento denominado PAI, sí sienta las bases conceptuales de la planificación individualizada de la atención.

La concreción del PAI se desarrolla principalmente en el ámbito autonómico, donde se establecen requisitos formales sobre su contenido, periodicidad de revisión y responsabilidades profesionales. En la Comunidad de Madrid, la Orden de 14 de noviembre de 2024, publicada en el Boletín Oficial de la Comunidad de Madrid (BOCM) [20], regula de manera explícita la elaboración, seguimiento y actualización del Plan de Atención Individualizada en centros residenciales para personas mayores, consolidándolo como herramienta obligatoria de planificación y evaluación de la atención integral.

La elaboración del PAI exige la participación coordinada de múltiples perfiles profesionales, entre los que se incluyen médicos, personal de enfermería, fisioterapeutas, terapeutas ocupacionales, psicólogos, trabajadores sociales y otros profesionales del ámbito asistencial. Cada disciplina aporta valoraciones específicas basadas en pruebas estandarizadas, observaciones clínicas y registros evolutivos, lo que convierte al PAI en un documento interdisciplinar por naturaleza.

Entre las herramientas utilizadas para su elaboración destacan escalas funcionales, cognitivas, emocionales y de riesgo, destinadas a evaluar aspectos como la movilidad, el equilibrio, el dolor, el estado nutricional, la integridad cutánea o el deterioro cognitivo. Estas valoraciones permiten establecer objetivos individualizados, planificar intervenciones y realizar un seguimiento evolutivo que facilite la detección precoz de cambios significativos en la situación del residente. En este sentido, el PAI no es únicamente un requisito documental, sino un mecanismo estructurado de planificación clínica y organizativa.

### **3.3 SISTEMAS DE GESTIÓN SOCIO SANITARIA ACTUALES**

La gestión de la atención sociosanitaria en centros residenciales ha experimentado una transformación progresiva en paralelo a la evolución del modelo asistencial. En sus primeras

etapas, la información clínica y asistencial se registraba mayoritariamente en soporte papel, mediante documentos independientes que dificultaban la visión global del residente y el seguimiento longitudinal de su evolución.

El incremento de la complejidad asistencial, asociado al envejecimiento poblacional y a la elevada prevalencia de pluripatologías y situaciones de dependencia, evidenció la necesidad de mejorar la organización, trazabilidad y accesibilidad de la información. Esta situación impulsó la adopción de sistemas informatizados que, en una fase inicial, se orientaron principalmente a la gestión administrativa y económica de los centros, incluyendo facturación, control de plazas y gestión de recursos humanos.

Posteriormente, estos sistemas evolucionaron hacia plataformas integrales de gestión sociosanitaria capaces de centralizar la información clínica, funcional y social del residente. Los actuales sistemas de información, frecuentemente estructurados en forma de ERP especializados, permiten registrar de manera estandarizada las valoraciones realizadas por los distintos profesionales, almacenar históricos completos de escalas y pruebas, y documentar intervenciones, incidencias y observaciones diarias.

En los últimos años, la digitalización del sector ha avanzado hacia modelos más conectados, impulsados por la necesidad de interoperabilidad entre sistemas sanitarios y sociales. La integración con historias clínicas electrónicas, plataformas regionales de salud o sistemas de receta electrónica constituye una tendencia creciente, especialmente en entornos donde se busca garantizar la continuidad asistencial entre atención primaria, hospitalaria y residencial. No obstante, el grado de integración varía significativamente entre territorios y proveedores tecnológicos.

De forma paralela, comienza a observarse un interés creciente por la analítica de datos aplicada al ámbito sociosanitario. El uso de datalakes corporativos, cuadros de mando y herramientas de business intelligence permite explotar grandes volúmenes de información histórica con fines organizativos, de control de calidad o de apoyo a la toma de decisiones. Sin embargo, esta explotación suele centrarse en indicadores agregados o métricas

operativas, y no tanto en la síntesis automatizada de información individualizada para la elaboración de documentos clínicos complejos.

En este contexto, aunque los sistemas actuales ofrecen una capacidad sólida de registro, almacenamiento y consulta de información asistencial, su funcionalidad se orienta principalmente a la gestión y estructuración de datos, sin incorporar mecanismos avanzados de síntesis o interpretación automatizada. Esta situación pone de manifiesto la existencia de un margen de mejora en la explotación inteligente de la información registrada.

### ***3.4 AUTOMATIZACIÓN Y GENERACIÓN DE INFORMES EN EL ÁMBITO CLÍNICO***

La automatización documental en el ámbito clínico no es un fenómeno reciente. Desde la implantación de historias clínicas electrónicas, numerosos sistemas han incorporado funcionalidades orientadas a la generación automática de informes a partir de datos estructurados previamente registrados. Estos mecanismos suelen basarse en plantillas parametrizadas que permiten insertar resultados de pruebas, diagnósticos codificados o tratamientos prescritos en documentos estandarizados.

En el entorno hospitalario, es habitual la generación automática de informes de alta, resultados de laboratorio o pruebas diagnósticas, donde la estructura del documento está altamente definida y el contenido procede directamente de campos estructurados. Este tipo de automatización se basa fundamentalmente en reglas predefinidas y ensamblado de datos, sin realizar una interpretación clínica compleja.

En el ámbito sociosanitario, sin embargo, la situación es diferente. Aunque los sistemas de gestión permiten exportar información y generar determinados informes básicos, la elaboración del Plan de Atención Individualizada y de sus seguimientos requiere integrar información procedente de múltiples escalas, notas evolutivas y observaciones interdisciplinares. Esta tarea implica un proceso de síntesis que, en la mayoría de los centros, continúa realizándose de forma manual.

Por tanto, aunque existen mecanismos de automatización documental, estos suelen limitarse a la estructuración de datos y no a la interpretación evolutiva de la información. Esta limitación constituye uno de los principales retos que motivan la incorporación controlada de técnicas de Inteligencia Artificial en la generación de informes clínicos complejos.

### ***3.5 USO DE INTELIGENCIA ARTIFICIAL EN LA SÍNTESIS DE INFORMACIÓN CLÍNICA***

El avance de las técnicas de Inteligencia Artificial, y en particular del Procesamiento del Lenguaje Natural (NLP), ha abierto nuevas posibilidades para la síntesis automática de información clínica y asistencial. Estas tecnologías permiten analizar grandes volúmenes de texto no estructurado, como notas evolutivas, observaciones profesionales o registros interdisciplinarios, y generar resúmenes coherentes que facilitan la interpretación global de la información registrada.

Los Modelos de Lenguaje de Gran Tamaño (LLM) han demostrado una elevada capacidad para generar texto contextualizado, reformular contenidos y sintetizar documentación compleja. En el ámbito sanitario, estas capacidades resultan especialmente relevantes en tareas como la redacción asistida de informes clínicos, la estructuración de historias médicas o la extracción de información clave a partir de registros extensos. Su potencial radica en la posibilidad de transformar datos dispersos en narrativas comprensibles y estructuradas.

Sin embargo, la aplicación de estos modelos en entornos asistenciales plantea retos significativos. Entre ellos destacan el riesgo de generación de información no fundamentada (hallucinations) [21], la limitada explicabilidad de los procesos internos del modelo, la posible introducción de sesgos y la necesidad de garantizar la trazabilidad de las fuentes utilizadas. En contextos como el sociosanitario, donde las decisiones afectan directamente a la salud y bienestar de personas vulnerables, estos riesgos adquieren una especial relevancia.

Para mitigar estas limitaciones, se han desarrollado enfoques como la Generación Aumentada por Recuperación (RAG), que combinan modelos generativos con mecanismos

de recuperación de información procedente de fuentes verificadas. Este enfoque permite restringir la generación de contenido a datos reales previamente almacenados en los sistemas corporativos, reduciendo la probabilidad de inferencias no justificadas y aumentando la fiabilidad del resultado final.

Además, el marco regulatorio europeo reciente, especialmente el Reglamento de Inteligencia Artificial de la Unión Europea (AI Act) [22], refuerza la necesidad de aplicar estas tecnologías bajo criterios de transparencia, supervisión humana y gestión del riesgo. En consecuencia, la integración de la IA en el ámbito sociosanitario debe plantearse desde un enfoque asistido y controlado, en el que el profesional mantiene en todo momento la responsabilidad sobre la validación y la toma de decisiones.

A pesar del potencial demostrado por estas tecnologías, la adopción de soluciones basadas en Inteligencia Artificial para la generación de informes en centros residenciales sigue siendo limitada. La combinación de exigencias normativas, sensibilidad de los datos y complejidad organizativa hace necesario el desarrollo de sistemas específicos que integren la IA como herramienta de apoyo a la labor profesional, y no como sustituto de su criterio clínico.

## **Capítulo 4. DEFINICIÓN DEL TRABAJO**

En este capítulo se define el alcance del proyecto desarrollado, estableciendo su justificación técnica y funcional a partir del análisis del estado de la cuestión presentado en capítulos anteriores. Asimismo, se detallan los objetivos generales y específicos que guían el desarrollo del sistema propuesto, junto con la metodología empleada para su consecución y la planificación temporal y económica asociada. Este capítulo permite contextualizar el proyecto desde una perspectiva técnica y organizativa, delimitando claramente qué se pretende alcanzar y en qué condiciones se llevará a cabo su desarrollo.

### **4.1 JUSTIFICACIÓN**

A pesar de los avances tecnológicos y organizativos en el ámbito sociosanitario expuestos en el capítulo anterior, el uso de los sistemas de información actuales continúa orientándose mayoritariamente al registro, almacenamiento y consulta de datos, mientras que la explotación avanzada de la información para la elaboración de informes de seguimiento sigue siendo, en muchos casos, un proceso manual y costoso en términos de tiempo y recursos.

En el contexto de los centros residenciales, la elaboración de informes de seguimiento asociados al Plan de Atención Individualizada (PAI) requiere que los profesionales revisen extensos historiales de notas clínicas, escalas de valoración y registros asistenciales dispersos en el sistema de gestión. Esta tarea de síntesis documental supone una carga administrativa significativa, que reduce el tiempo disponible para la atención directa al residente y aumenta el riesgo de omisiones, inconsistencias o falta de homogeneidad en los informes generados.

Esta situación pone de manifiesto la necesidad de evolucionar desde sistemas centrados exclusivamente en el almacenamiento de información hacia soluciones capaces de transformar los datos registrados en conocimiento estructurado y útil, facilitando la toma de

decisiones clínicas y mejorando la eficiencia de los procesos asistenciales. En este sentido, la integración de técnicas de Inteligencia Artificial aplicadas de forma controlada, junto con arquitecturas modernas de datos, abre la puerta a nuevos modelos de apoyo a la labor profesional.

El presente proyecto se justifica, por tanto, en la necesidad real del sector sociosanitario de disponer de herramientas que asistan a los profesionales en la generación de informes de seguimiento, sin sustituir su criterio clínico, pero reduciendo la carga administrativa asociada a tareas repetitivas de síntesis documental. Desde un punto de vista técnico y de mercado, la solución propuesta responde a una problemática actual y generalizada en los centros residenciales, aportando un valor diferencial mediante la automatización parcial y segura de la redacción del PAI a partir de los datos ya existentes en los sistemas corporativos.

## **4.2 OBJETIVOS**

El objetivo final de este Trabajo de Fin de Máster es desarrollar un sistema de generación asistida y automatizada de informes de seguimiento del Plan de Atención Individualizada (PAI) en centros residenciales, que permita transformar los datos registrados en el sistema de gestión y en el datalake corporativo en borradores estructurados, reduciendo la carga administrativa y mejorando la eficiencia del proceso de elaboración de informes. Para la consecución de este es necesario cumplir las siguientes metas:

- Analizar la estructura y tipología de los datos clínicos y asistenciales registrados en el sistema de gestión sociosanitaria. Este objetivo implica estudiar el modelo de datos existente, identificar las tablas y campos relevantes para la elaboración del PAI y comprender la naturaleza de la información almacenada, tanto estructurada (escalas, resultados numéricos, fechas) como no estructurada (observaciones clínicas). Este análisis es fundamental para determinar qué datos pueden ser explotados y cómo deben ser tratados.
- Diseñar un proceso de extracción y transformación de datos desde el datalake. Se pretende definir un flujo técnico que permita acceder de forma eficiente a los datos

almacenados en el datalake corporativo, filtrarlos por residente y residencia, y transformarlos en estructuras adecuadas para su explotación posterior. Este proceso incluye la normalización de campos, la gestión de formatos temporales y la preparación de conjuntos de datos listos para su integración en el informe.

- Implementar mecanismos de síntesis de información clínica mediante técnicas de Procesamiento del Lenguaje Natural. Este objetivo consiste en integrar modelos de lenguaje que permitan generar resúmenes interpretativos a partir de datos históricos y observaciones clínicas, especialmente en aquellos apartados donde se requiere análisis evolutivo. La implementación debe realizarse bajo criterios de control y limitación de inferencias, garantizando que el contenido generado se base exclusivamente en los datos disponibles.
- Definir y validar prompts orientados a la generación estructurada de contenido. Se busca diseñar instrucciones específicas para el modelo de lenguaje que delimiten el estilo, la extensión, el enfoque interpretativo y las restricciones semánticas del texto generado. Asimismo, se debe validar que dichos prompts producen resultados coherentes, ajustados al contexto sociosanitario y alineados con el lenguaje utilizado en los informes dirigidos a residentes y familiares.
- Desarrollar un sistema de generación de informes basado en plantillas editables. Este objetivo implica la creación de una arquitectura documental que combine una plantilla estructurada en formato Word con un sistema de sustitución dinámica de etiquetas. Se persigue garantizar uniformidad en la estructura de los informes, facilitar su mantenimiento y permitir la inserción automática de texto, tablas, gráficos y bloques generados mediante IA.
- Establecer mecanismos de control y validación para minimizar errores. Se pretende incorporar procedimientos que eviten la generación de contenido no verificado, la permanencia de etiquetas sin reemplazar o la inserción de datos inexistentes. Esto incluye la gestión explícita de excepciones documentales, la verificación de disponibilidad de datos y la diferenciación visual de posibles incidencias en el documento generado.

- Evaluar los resultados obtenidos desde el punto de vista funcional y asistencial. Finalmente, se busca analizar el impacto del sistema desarrollado en términos de funcionamiento técnico, reducción del tiempo de elaboración de informes y percepción por parte de los profesionales. Esta evaluación permitirá determinar el grado de cumplimiento de los objetivos planteados y valorar la viabilidad de su implantación en un entorno real.

### **4.3 METODOLOGÍA**

Para el desarrollo de este proyecto se ha seguido una metodología de trabajo Agile [23]. De esta manera se han agrupado las tareas por bloques y se han adaptado las fechas en función del progreso y los imprevistos.

En una primera fase, se ha llevado a cabo un análisis del contexto asistencial y de los flujos de información existentes, mediante reuniones con los distintos perfiles profesionales implicados en la elaboración y uso de los informes PAI. Estas sesiones han permitido identificar las principales dificultades del proceso actual, comprender el lenguaje y la estructura de los informes, y recoger los requisitos funcionales y no funcionales del sistema.

A continuación, se ha realizado un análisis técnico de los sistemas de información disponibles, centrado en el ERP ResiPlus y en el datalake corporativo, con el objetivo de identificar las fuentes de datos relevantes, su estructura y su potencial de explotación. Sobre esta base, se ha diseñado posteriormente el flujo de extracción, tratamiento y preparación de la información necesario para alimentar los módulos de generación de informes.

En paralelo, se ha abordado el diseño del modelo de Inteligencia Artificial, adoptando un enfoque de uso controlado basado en la definición cuidadosa de prompts y en la limitación estricta del contexto de actuación del modelo. Este proceso se ha desarrollado de forma iterativa, ajustando las instrucciones a partir de pruebas sucesivas y validaciones funcionales, con el objetivo de garantizar que los textos generados fueran coherentes, comprensibles y clínicamente seguros.

Finalmente, se ha implementado el sistema completo, integrando los distintos módulos desarrollados (extracción de datos, tratamiento de la información, generación de plantillas y síntesis asistida mediante IA) y se han llevado a cabo pruebas de validación en escenarios reales. La metodología se completa con una fase de evaluación de resultados y recogida de feedback por parte de los profesionales, que permite valorar el impacto real de la herramienta y detectar posibles líneas de mejora.

#### ***4.4 PLANIFICACIÓN Y ESTIMACIÓN ECONÓMICA***

En este apartado se presenta la planificación temporal de las actividades necesarias para el desarrollo del proyecto, así como una estimación orientativa del esfuerzo y los costes asociados. La planificación se ha diseñado utilizando un diagrama de Gantt orientativo (Figura 4) [24] para asegurar una ejecución ordenada del proyecto, contemplando márgenes de ajuste ante posibles desviaciones técnicas o funcionales. Por ejemplo, se ha conseguido completar de manera mucho más ágil la parte de desarrollo de la herramienta, mientras que las primeras semanas de entender el problema han requerido más tiempo de lo esperado.

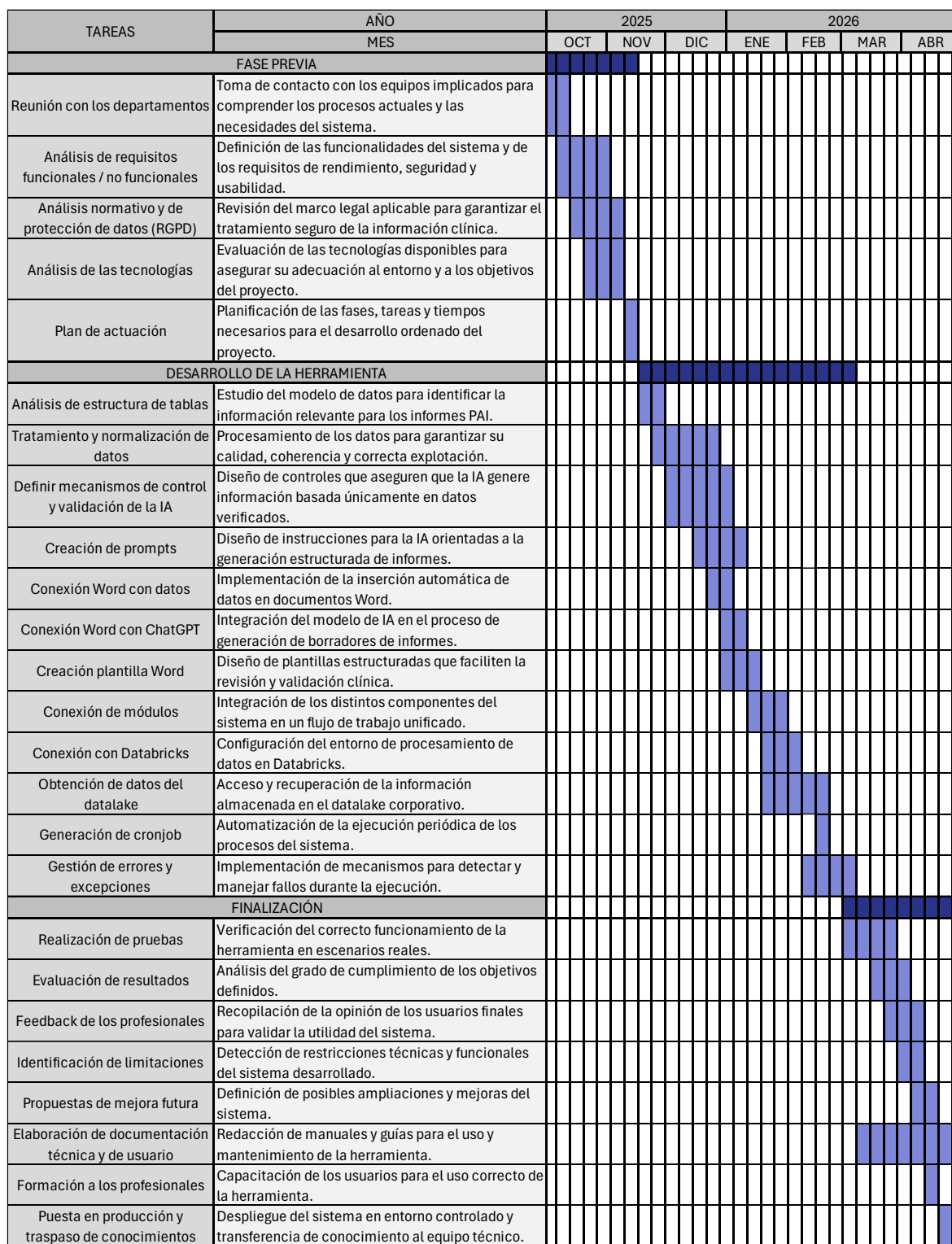


Figura 4 - Diagrama de Gantt del proyecto

En cuanto a la estimación económica, dado que el proyecto se ha desarrollado utilizando infraestructuras ya disponibles en la organización, los costes se han centrado fundamentalmente en el esfuerzo humano dedicado al desarrollo. A efectos orientativos, se ha considerado un perfil técnico con conocimientos en ingeniería de datos y desarrollo de software, con una dedicación parcial a lo largo del periodo del proyecto.

*Tabla 1 - Costes asociados al equipo que desarrolla el proyecto*

| <i>Perfil</i>                           | <i>Dedicación (horas)</i> | <i>Coste (€/hora)</i> | <i>Coste total (€)</i> |
|-----------------------------------------|---------------------------|-----------------------|------------------------|
| Desarrollador principal                 | 450                       | 30                    | 13.500                 |
| Profesional sociosanitario (validación) | 20                        | 25                    | 500                    |
| Técnico de sistemas / datos             | 15                        | 35                    | 525                    |
| Responsable funcional                   | 10                        | 40                    | 400                    |
| <b>Total estimado</b>                   |                           |                       | <b>14.925</b>          |

Adicionalmente, se ha tenido en cuenta el coste asociado al uso de servicios en la nube y modelos de Inteligencia Artificial, si bien estos se han mantenido en un entorno controlado y de pruebas, minimizando el impacto económico. En un escenario de despliegue real a mayor escala, estos costes deberían ser evaluados en función del volumen de datos, la frecuencia de ejecución y el número de centros usuarios.

Tabla 2 - Costes asociados a los servicios utilizados

| <i>Recurso</i>                | <i>Coste mensual estimado (€)</i> |
|-------------------------------|-----------------------------------|
| Azure Data Lake Storage       | 20                                |
| Databricks (uso intermitente) | 50                                |
| SharePoint / Excel inputs     | 0 (incluido)                      |
| Otros servicios auxiliares    | 10                                |
| <b>Total</b>                  | <b>80</b>                         |

Finalmente, se ha estimado el coste asociado al uso de los modelos de Inteligencia Artificial. Esto se ha realizado considerando un escenario conservador de generación de informes de seguimiento, con un número limitado de ejecuciones y uso controlado del modelo. Dado que la IA se emplea únicamente como apoyo a la redacción de borradores, el coste anual resulta reducido en comparación con el resto de los componentes del sistema.

Tabla 3 - Costes por uso de los modelos de Inteligencia Artificial

| <i>Concepto</i>            | <i>Coste estimado (€)</i> |
|----------------------------|---------------------------|
| Coste por informe generado | 0,10                      |
| Informes anuales estimados | 520                       |
| <b>Coste anual IA</b>      | <b>52</b>                 |

A partir de las estimaciones anteriores, se ha realizado un análisis económico del proyecto diferenciando entre la fase inicial de desarrollo (Año 0) y los costes recurrentes de operación y mantenimiento en años sucesivos. El Año 0 comprende el periodo completo de desarrollo del sistema, desde octubre hasta abril, e incluye tanto las tareas de análisis y diseño como las de implementación técnica y validación.

Durante los primeros meses del proyecto, el esfuerzo se ha centrado principalmente en la comprensión del dominio sociosanitario, la revisión normativa, la definición funcional del sistema y la recopilación de requisitos mediante reuniones con profesionales. En esta fase inicial no fue necesario el uso de infraestructuras en la nube ni de modelos de Inteligencia Artificial, ya que el trabajo se orientó a tareas conceptuales y de diseño. El uso de servicios cloud y de IA se concentró, por tanto, en los meses finales del desarrollo, una vez definida la arquitectura y comenzada la implementación técnica.

*Tabla 4 - Costes de mantenimiento y explotación Año 0 + 5 primeros años*

| <i>Concepto</i>                 | <i>Año 0 (€)</i> | <i>Año 1 (€)</i> | <i>Año 2 (€)</i> | <i>Año 3 (€)</i> | <i>Año 4 (€)</i> | <i>Año 5 (€)</i> |
|---------------------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| <b>Gasto de personal</b>        |                  |                  |                  |                  |                  |                  |
| Horas anuales<br>(horas)        | 450              | 104              | 104              | 104              | 104              | 104              |
| Coste horario<br>medio (€/hora) | 30               | 30               | 30               | 30               | 30               | 30               |
| Total gasto de<br>personal      | 13.500           | 3.120            | 3.120            | 3.120            | 3.120            | 3.120            |

| Concepto                                 | Año 0 (€)     | Año 1 (€)    | Año 2 (€)    | Año 3 (€)    | Año 4 (€)    | Año 5 (€)    |
|------------------------------------------|---------------|--------------|--------------|--------------|--------------|--------------|
| <b>Infraestructura y servicios cloud</b> |               |              |              |              |              |              |
| Azure Data Lake Storage                  | 60            | 240          | 240          | 240          | 240          | 240          |
| Databricks                               | -             | -            | 600          | 600          | 600          | 600          |
| Otros servicios auxiliares               | -             | 120          | 120          | 120          | 120          | 120          |
| Total infraestructura cloud              | 60            | 360          | 960          | 960          | 960          | 960          |
| <b>Uso de IA</b>                         |               |              |              |              |              |              |
| Coste generación de informes             | 20            | 200          | 200          | 200          | 200          | 200          |
| <b>Costes administrativos y soporte</b>  |               |              |              |              |              |              |
| Soporte técnico puntual                  | -             | 175          | 175          | 175          | 175          | 175          |
| <b>Total coste anual</b>                 | <b>13.580</b> | <b>3.855</b> | <b>4.455</b> | <b>4.455</b> | <b>4.455</b> | <b>4.455</b> |

El análisis económico muestra que el principal esfuerzo económico se concentra en el Año 0, correspondiente al desarrollo del sistema, donde predominan los costes asociados al trabajo técnico y a las tareas de análisis, diseño e implementación. En esta fase, el uso de infraestructura cloud y modelos de Inteligencia Artificial es limitado, ya que su utilización se concentra en los meses finales del proyecto.

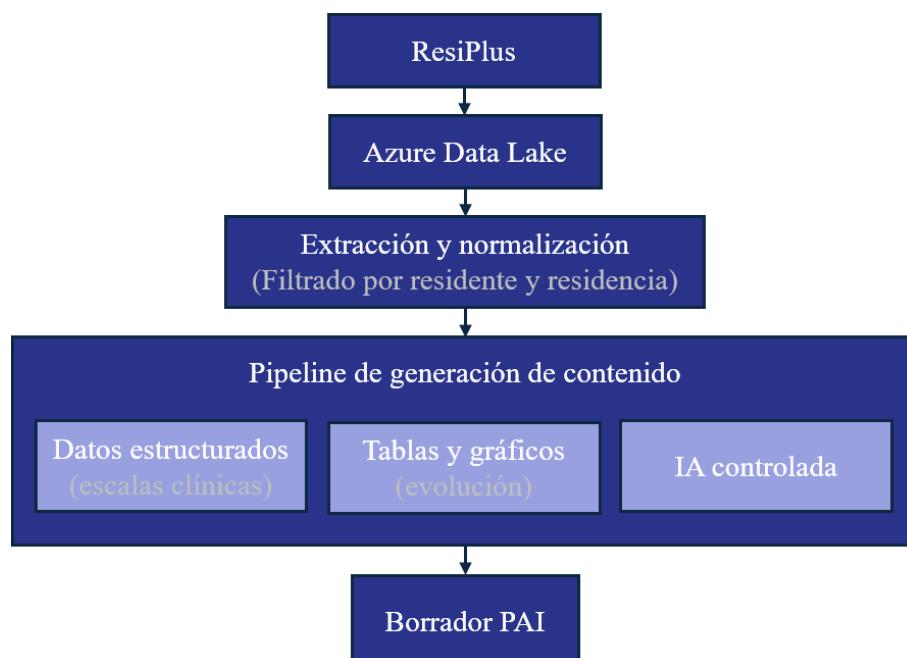
En los años posteriores, el sistema presenta costes recurrentes bajos y estables, asociados principalmente al mantenimiento técnico, al uso moderado de servicios en la nube y a la generación controlada de informes mediante IA. Esta estructura de costes favorece la sostenibilidad y escalabilidad del sistema, permitiendo su despliegue progresivo en distintos centros residenciales con un impacto económico reducido.

Cabe destacar que, aunque el sistema no genera un ahorro salarial directo, el valor económico del proyecto se manifiesta en términos de eficiencia operativa, mejora de la calidad asistencial, reducción de carga administrativa y mayor trazabilidad de la información clínica, aspectos clave en entornos sociosanitarios.

## Capítulo 5. SISTEMA DESARROLLADO

En este bloque se describe de forma detallada el sistema desarrollado para la generación automatizada de informes de seguimiento del Plan de Atención Individualizada (PAI) en centros residenciales. Tras haber definido el contexto, los objetivos y la metodología del proyecto en los capítulos anteriores, a continuación se aborda la implementación técnica de la solución propuesta.

Dado que el sistema integra distintos componentes (obtención y tratamiento de datos, arquitectura del programa, generación documental y uso controlado de Inteligencia Artificial) se ha optado por estructurar su descripción en varios capítulos diferenciados, cada uno centrado en una parte específica del desarrollo. El esquema del sistema desarrollado se puede ver en la Figura 5.



*Figura 5 - Arquitectura del sistema*

El Capítulo 6 aborda las fuentes de datos utilizadas y los procesos de extracción, tratamiento y normalización de la información. El Capítulo 7 describe la arquitectura general del programa y la organización interna del código. En el Capítulo 8 se detalla el sistema de generación de informes y plantillas documentales, incluyendo el mecanismo de sustitución dinámica de contenido. El Capítulo 9 presenta el modelo controlado de Inteligencia Artificial, explicando los principios de diseño, la estructuración de prompts y las medidas de control implementadas.

## **Capítulo 6. DATOS DE ENTRADA Y TRATAMIENTO DE LA INFORMACIÓN**

En este capítulo se describen las fuentes de información utilizadas por el sistema, así como los procesos realizados para la obtención, tratamiento y preparación de los datos empleados antes de su explotación.

### **6.1 ORIGEN DE LOS DATOS**

Los datos utilizados en este proyecto han sido extraídos de dos fuentes principales: el sistema de gestión residencial ResiPlus y el datalake corporativo de la empresa. Estos sistemas centralizan toda la información clínica, funcional y asistencial registrada en cada una de las residencias.

En primer lugar, tal y como se ha descrito de forma general en capítulos anteriores, ResiPlus es el sistema de gestión utilizado en los centros residenciales para el registro de la actividad asistencial diaria. Este sistema centraliza la información de cada residente mediante un modelo de datos relacional, en el que la información se organiza en distintas tablas asociadas a entidades como residentes, valoraciones, escalas, evolutivos o intervenciones. Cada residente se identifica mediante un identificador único (IDResidente), que permite relacionar toda la información registrada a lo largo del tiempo. A partir de esta entidad principal, se estructuran diferentes módulos funcionales que almacenan valoraciones medicas mediante escalas clínicas, registros evolutivos y anotaciones de los profesionales, información administrativa y de estancia en el centro o datos clínicos y asistenciales asociados a la atención diaria.

Esta organización permite mantener un histórico completo de la evolución del residente, aunque la información se encuentra distribuida en múltiples tablas y módulos, lo que dificulta su explotación directa para la elaboración de informes.

Por otro lado, el datalake corporativo actúa como repositorio centralizado donde se integran los datos procedentes de ResiPlus y de otros sistemas de información de la organización. En este entorno, los datos se almacenan generalmente en formatos optimizados para su procesamiento analítico, como archivos Parquet, lo que permite trabajar con grandes volúmenes de información de forma eficiente. El datalake facilita el acceso unificado a la información histórica y permite aplicar procesos de transformación y preparación de datos previos a su explotación por los distintos componentes del sistema desarrollado.



*Figura 6 - Proceso de extracción y tratamiento de la información*

## **6.2 TIPOLOGÍA DE LOS DATOS TRATADOS**

Los datos utilizados en el proyecto presentan una naturaleza heterogénea, pudiendo clasificarse en tres grandes categorías: datos estructurados, datos no estructurados y datos temporales. Esta clasificación resulta relevante, ya que condiciona tanto los procesos de tratamiento como las técnicas utilizadas para su explotación.

### **6.2.1 DATOS ESTRUCTURADOS: ESCALAS CLÍNICAS Y REGISTROS CUANTITATIVOS**

Los datos estructurados corresponden a información almacenada en campos definidos dentro de tablas, con formatos bien establecidos y fácilmente procesables mediante consultas SQL. Dentro de esta categoría se incluyen principalmente resultados de escalas clínicas como Barthel, Tinetti o Norton, valoraciones numéricas asociadas a pruebas, fechas de realización de las evaluaciones o identificadores de residente y residencia.

Estas escalas permiten cuantificar distintos aspectos del estado del residente, como la funcionalidad, el riesgo de caídas, el estado nutricional o el deterioro cognitivo. Su carácter estructurado facilita su explotación para generar tablas, gráficos de evolución y resúmenes cuantitativos dentro de los informes.

*Tabla 5 - Ejemplo datos estructurados*

| <i>Fecha</i> | <i>Resultado</i> |
|--------------|------------------|
| 05/09/2022   | 16               |
| 10/03/2023   | 16               |
| 12/09/2024   | 11               |
| 04/04/2025   | 12               |
| 02/10/2025   | 12               |

### **6.2.2 DATOS NO ESTRUCTURADOS: NOTAS Y OBSERVACIONES CLÍNICAS**

Los datos no estructurados corresponden principalmente a texto libre introducido por los profesionales en el sistema de gestión. Incluyen notas evolutivas, observaciones clínicas, comentarios asistenciales o descripciones cualitativas del estado del residente en ese momento.

Este tipo de información resulta especialmente relevante, ya que aporta contexto clínico y permite interpretar la evolución del residente más allá de los valores numéricos. Sin embargo, su tratamiento es más complejo, ya que no sigue una estructura fija y requiere el uso de técnicas de Procesamiento del Lenguaje Natural (NLP) para su análisis y síntesis.

|                                                                                                                           |                                                                                                                                                                                                                                                      |
|---------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Residente femenina entorno a los 90 años de edad, institucionalizada en nuestro centro desde el 2021. Hija muy pendiente. | Tiene dependencia leve para las ABVD (Barthel 75/100): independiente para comer, ducha, aseo personal, traslados y desplazamientos; necesita ayuda para vestirse, uso de retrete y subir y bajar escalones. Incontinente urinaria, continente fecal. |
| Actualmente presenta una adecuada ingesta oral con una dieta basal de consistencia normal. Tiene normopeso.               |                                                                                                                                                                                                                                                      |
| Deterioro cognitivo leve (FAST 4/7f, MEC 23/35), no estudiado.                                                            | Utiliza andador como aparato de movilidad asistida. Riesgo medio de caída (Tinetti 15/28; Equilibrio 6/16, marcha 9/12).                                                                                                                             |

*Figura 7 - Ejemplo datos no estructurados*

### 6.2.3 DATOS TEMPORALES: REGISTROS EVOLUTIVOS

Una característica fundamental de los datos utilizados es su componente temporal. Tanto las escalas como las observaciones se registran en distintos momentos del tiempo, lo que permite analizar la evolución del residente. Estos datos incluyen fechas de realización de escalas, históricos de resultados o secuencias de anotaciones clínicas.

El tratamiento adecuado de esta dimensión temporal es clave para poder generar informes de seguimiento que reflejen tendencias, cambios significativos y evolución clínica del residente.



*Figura 8 - Ejemplo evolutivo Norton*

### **6.3 PROCESOS DE EXTRACCIÓN DE DATOS**

El acceso a los datos utilizados en el proyecto se realiza directamente desde el datalake corporativo, donde la información procedente del sistema de gestión ResiPlus se encuentra previamente integrada y almacenada en formatos optimizados para su explotación analítica.

El datalake actúa como capa intermedia entre los sistemas operacionales y los procesos de análisis, permitiendo disponer de datos históricos consolidados, estructurados y accesibles de forma eficiente. En este entorno, la información se organiza en distintos conjuntos de datos que reflejan las entidades principales del sistema, como residentes, escalas, valoraciones y registros asistenciales.

El proceso de extracción de datos se basa en la selección de la información relevante para cada caso de uso, aplicando criterios de filtrado específicos. De forma general, este proceso incluye:

1. Identificación de las fuentes de datos relevantes dentro del datalake, en función del tipo de información requerida (escalas, observaciones, etc.).
2. Selección de campos necesarios para la generación de informes, evitando el uso de información redundante o no relevante.
3. Filtrado por residente y residencia, utilizando identificadores internos que permiten aislar la información correspondiente a cada caso.
4. Ordenación temporal de los registros, con el objetivo de preservar la secuencia evolutiva de la información.

Este enfoque permite obtener conjuntos de datos específicos, coherentes y listos para su posterior tratamiento, sin necesidad de acceder directamente al sistema transaccional (ResiPlus), lo que reduce la carga sobre el sistema operativo y mejora la eficiencia del procesamiento.

Además, el uso del datalake facilita la escalabilidad del sistema, permitiendo trabajar con grandes volúmenes de información histórica y garantizando la disponibilidad de los datos para su explotación por los distintos módulos del sistema desarrollado.

## ***6.4 TRATAMIENTO, NORMALIZACIÓN Y ANONIMIZACIÓN DE LA INFORMACIÓN***

Una vez extraídos, los datos son sometidos a un conjunto de procesos de transformación orientados a garantizar su calidad, coherencia y adecuación para su uso en el sistema.

En primer lugar, se realizan tareas de normalización, como la conversión de formatos de fecha, la homogeneización de nombres de columnas, la limpieza de texto en campos de observaciones (eliminando saltos de línea o caracteres no deseados) o la validación de la existencia de las columnas necesarias para cada uno de los procesos. Asimismo, se implementan mecanismos de control de calidad de datos, como la detección de valores nulos o inconsistentes, con el objetivo de evitar errores en la generación de los informes.

En relación con la seguridad y el cumplimiento normativo, el tratamiento de la información se ha realizado siguiendo los principios establecidos en el Reglamento General de Protección de Datos (RGPD). En particular, se ha trabajado con identificadores internos del sistema en lugar de datos personales directos, minimizando la exposición de información sensible durante el procesamiento.

Adicionalmente, el sistema incorpora mecanismos de gestión de excepciones que permiten detectar situaciones como la ausencia de datos, la falta de registros para determinadas escalas o la presencia de información incompleta. En estos casos, se generan mensajes controlados que se incorporan al informe, evitando resultados incorrectos o interpretaciones erróneas.

## **Capítulo 7. ARQUITECTURA Y ESTRUCTURA DEL PROGRAMA**

El sistema desarrollado para la generación automatizada de informes del Plan de Atención Individualizada (PAI) se basa en una arquitectura modular orientada a la separación de responsabilidades, la reutilización de componentes y el control del flujo de ejecución.

A diferencia de un diseño puramente teórico por capas, la arquitectura implementada se articula en torno a una organización práctica del código, estructurada en módulos funcionales que reflejan directamente las distintas etapas del proceso de generación del informe. Este enfoque permite una correspondencia clara entre la lógica del sistema y su implementación, facilitando tanto el mantenimiento como la escalabilidad.

### ***7.1 VISIÓN GENERAL DE LA ARQUITECTURA***

El sistema sigue un modelo de procesamiento secuencial organizado como un pipeline de generación documental. En este modelo, la información fluye desde las fuentes de datos hasta el documento final a través de una serie de etapas claramente definidas.

Cada etapa del proceso se encapsula en módulos independientes que interactúan entre sí mediante estructuras de datos bien definidas, evitando acoplamientos innecesarios.

El sistema no delega la generación completa del informe en la Inteligencia Artificial, sino que adopta un enfoque híbrido que combina datos estructurados procedentes del datalake, procesamiento programático, generación controlada de texto mediante IA e inserción dinámica del contenido en una plantilla documental. Esta combinación permite mantener el control sobre el resultado generado y garantizar la coherencia del informe final.

## 7.2 *FLUJO DE EJECUCIÓN DEL SISTEMA*

El flujo de ejecución del sistema puede describirse como una secuencia ordenada de operaciones:

1. Carga de los datos del residente seleccionado desde el datalake corporativo.
2. Procesamiento y estructuración de la información en formatos internos (principalmente DataFrames).
3. Ejecución del pipeline, que coordina la generación de cada sección del informe borrador.
4. Generación de contenido específico, incluyendo tablas, gráficos y bloques interpretativos.
5. Construcción de un diccionario de etiquetas, donde cada marcador del documento se asocia a su contenido correspondiente.
6. Sustitución de etiquetas en la plantilla Word.
7. Generación del documento final, listo para revisión de los diferentes profesionales de las residencias.

Este flujo garantiza que cada componente del sistema actúe sobre información previamente validada, evitando dependencias circulares y facilitando la trazabilidad del proceso.

## 7.3 *ESTRUCTURA DEL PROYECTO*

El proyecto se organiza en distintos módulos que reflejan las responsabilidades principales del sistema. La siguiente tabla recoge los principales módulos auxiliares del sistema, junto con su función específica dentro del pipeline de generación documental.

Esta organización permite aislar la lógica de negocio, reducir la complejidad del código y facilitar la incorporación de nuevas funcionalidades sin afectar al núcleo del sistema. En los siguientes apartados se detalla cada uno de ellos.

Tabla 6 - Estructura del proyecto

| Módulo     | Función                                   |
|------------|-------------------------------------------|
| pipeline/  | Orquestación del proceso completo         |
| sections/  | Lógica específica por dominio asistencial |
| utils/     | Funciones comunes (Word, gráficos, IA)    |
| config/    | Configuración del sistema                 |
| templates/ | Plantillas documentales                   |

```

VALDELUZ-PAI /
├── resources/
│   ├── drivers/
│   ├── input/
│   │   └── ExcelInput.xlsx
│   └── templates/
│       └── Plantilla_PAI.docx
├── src/
│   ├── common/
│   │   ├── config/
│   │   ├── db/
│   │   ├── escalas/
│   │   ├── io/
│   │   ├── utils/
│   │   └── word/
│   ├── pipeline/
│   │   ├── build_context.py
│   │   ├── render_doc.py
│   │   └── run.py
│   ├── sections/
│   │   ├── animador/
│   │   ├── enfermeria/
│   │   ├── fisio/
│   │   ├── medico/
│   │   ├── psicologia/
│   │   ├── terapia_ocupacional/
│   │   └── trabajo_social/
│   └── temp/
├── config.py
├── main.py
└── openai_cost.py

```

## 7.4 *MÓDULO PIPELINE*

El pipeline actúa como controlador central del sistema, siendo responsable de invocar la carga de datos, ejecutar los distintos módulos de secciones, consolidar los resultados generados y construir la estructura final de etiquetas y contenidos. Este diseño permite modificar o ampliar el sistema añadiendo nuevas etapas sin alterar el resto de los módulos.

En la práctica, el pipeline se invoca mediante un script principal que recibe como parámetros de entrada el identificador del residente y el identificador de la residencia. A partir de estos parámetros, el sistema localiza y carga automáticamente la información correspondiente desde el datalake, inicializa los módulos de secciones activos y coordina su ejecución de forma secuencial. Una vez completada la generación de todos los bloques de contenido, el pipeline consolida los resultados en un diccionario de etiquetas y desencadena el proceso de construcción del documento final. Este diseño centralizado garantiza que cualquier modificación en el orden de ejecución o en los módulos activos pueda realizarse en un único punto del sistema, sin afectar al resto de los componentes.

## 7.5 *MÓDULO SECTIONS*

El módulo sections contiene la lógica específica asociada a cada área asistencial, como enfermería, fisioterapia o psicología. Cada sección se encarga de procesar los datos relevantes para su ámbito, generar tablas y estructuras de datos, crear gráficos cuando es necesario, invocar la generación de texto mediante IA en los casos requeridos y devolver los contenidos listos para su inserción en el documento.

Este enfoque permite encapsular el conocimiento funcional de cada dominio en módulos independientes, facilitando su mantenimiento y evolución. Además, la separación por áreas asistenciales permite añadir nuevas secciones sin modificar el resto del sistema, lo que aporta una alta escalabilidad.

## **7.6 *MÓDULOS AUXILIARES***

El sistema incorpora una serie de módulos auxiliares que proporcionan funcionalidades comunes reutilizables.

Los módulos de generación documental se encargan de sustituir las distintas etiquetas que se encuentran en la plantilla e insertar las tablas e imágenes necesarias. Los módulos de generación de gráficos construyen las visualizaciones a partir de datos históricos y las exportan en formatos compatibles con Word. Por su parte, los módulos de generación de texto mediante IA construyen los prompts estructurados, invocan el modelo de lenguaje y devuelven el texto controlado y validado. La centralización de estas funcionalidades evita la duplicación de código y mejora la consistencia del sistema.

## **7.7 *RECURSOS Y CONFIGURACIÓN EXTERNA***

El sistema hace uso de recursos externos que permiten desacoplar la lógica del programa de su configuración: las plantillas Word definen la estructura del informe, las claves de acceso a la API son necesarias para el uso del modelo de IA, y los parámetros de configuración recogen aspectos como rutas, formatos o modelos utilizados. Este diseño permite modificar aspectos del sistema sin necesidad de alterar el código fuente.

La gestión de las credenciales de acceso sigue un principio de separación estricta entre el código fuente y la configuración sensible. Las claves de API y los parámetros de conexión se almacenan en un fichero externo al repositorio, que no se incluye en el control de versiones, evitando así su exposición accidental. En el repositorio se incluye únicamente una plantilla de configuración con los campos necesarios pero sin valores reales, de modo que cualquier despliegue del sistema en un nuevo entorno requiera únicamente completar dicho fichero con las credenciales correspondientes.

## **7.8 ESCALABILIDAD, MANTENIMIENTO Y CONTROL DEL SISTEMA**

La arquitectura modular implementada permite la evolución del sistema de forma controlada. La incorporación de nuevas escalas clínicas, nuevas áreas asistenciales o cambios en la plantilla documental puede realizarse sin modificar la estructura global del programa.

El uso de un pipeline central y módulos independientes reduce el acoplamiento entre componentes, facilitando el mantenimiento y la detección de errores. A modo de ejemplo, la incorporación de una nueva área asistencial, como logopedia o terapia ocupacional, requeriría únicamente la creación de un nuevo módulo en el directorio sections/ siguiendo el patrón establecido por los módulos existentes, con sus correspondientes ficheros de procesamiento de datos y de generación de contenido mediante IA. El pipeline detectaría automáticamente el nuevo módulo y lo integraría en el flujo de ejecución sin necesidad de modificar ningún otro componente del sistema.

Esta capacidad de extensión sin modificación del núcleo del sistema es especialmente relevante en un entorno como el sociosanitario, donde los requisitos documentales pueden evolucionar en función de cambios normativos o de las necesidades específicas de cada centro residencial.

Además, el diseño del sistema permite rastrear el origen de cada elemento del informe generado, diferenciando claramente entre los datos estructurados procedentes del datalake, el contenido generado mediante procesamiento programático y los bloques producidos por la Inteligencia Artificial. Esta trazabilidad resulta fundamental en un entorno clínico, donde la fiabilidad de la información es un requisito crítico. Asimismo, el sistema mantiene al profesional como responsable final de la validación del informe, garantizando que la automatización no sustituya el criterio clínico.

## Capítulo 8. GENERACIÓN DE INFORMES Y

### PLANTILLAS

La generación automática de los informes de seguimiento del Plan de Atención Individualizada (PAI) constituye uno de los componentes centrales del sistema desarrollado. El enfoque adoptado no se basa en la redacción completa del documento mediante Inteligencia Artificial, sino en un modelo híbrido que combina una plantilla estructurada predefinida con la inserción dinámica de información procedente de bases de datos y bloques interpretativos generados de forma controlada.

Este planteamiento permite garantizar uniformidad documental, coherencia estructural y trazabilidad del contenido, minimizando los riesgos asociados a la generación automática de información en un entorno clínico. El sistema actúa como asistente en la elaboración del informe, manteniendo siempre al profesional como responsable final de su validación.

En este capítulo se describe el diseño del modelo documental, la justificación técnica de la estrategia adoptada y el mecanismo implementado para la sustitución dinámica de los distintos componentes del informe.

#### **8.1 DISEÑO DEL MODELO DE GENERACIÓN DOCUMENTAL**

El sistema de generación documental implementado en este proyecto se basa en el uso de una plantilla en formato Microsoft Word que contiene tanto texto fijo como un conjunto de etiquetas identificadas mediante el patrón <<NOMBRE\_ETIQUETA>>. Estas etiquetas actúan como marcadores que posteriormente serán reemplazados por información dinámica procedente de las bases de datos o generada mediante el modelo de lenguaje cuando se requiere interpretación clínica.

El funcionamiento del sistema comienza con la obtención de los datos necesarios desde el datalake corporativo, tal y como se ha descrito en el capítulo anterior. Una vez recuperada la información, esta es transformada y adaptada al formato requerido para su inserción en el documento. En esta fase se construye una estructura interna que asocia cada etiqueta de la plantilla con el contenido que debe sustituirla, pudiendo tratarse de texto simple, tablas estructuradas, gráficos previamente generados o bloques interpretativos producidos mediante Inteligencia Artificial bajo un conjunto de restricciones previamente definidas.

El programa recorre posteriormente el documento Word elemento por elemento, identificando la presencia de etiquetas tanto en párrafos como en tablas. Cuando se detecta una coincidencia, se procede a su reemplazo por el contenido correspondiente, manteniendo la coherencia de formato y respetando la estructura original de la plantilla. Este procedimiento permite desacoplar claramente la lógica de obtención de datos, la generación textual y la maquetación del documento, facilitando el mantenimiento del sistema y su posible ampliación futura.

El resultado final es un documento estructurado, uniforme y listo para revisión por parte del profesional, quien mantiene en todo momento el control y la validación última del contenido generado.

## **8.2 JUSTIFICACIÓN DEL USO DE PLANTILLAS ESTRUCTURADAS**

La decisión de utilizar una plantilla estructurada responde a criterios técnicos, organizativos y económicos. Desde el punto de vista documental, la plantilla garantiza uniformidad en todos los informes generados, asegurando que mantengan la misma estructura, orden de apartados y estilo formal. Esta homogeneidad facilita la lectura por parte de los profesionales, la comparación entre informes sucesivos y su correcta archivación dentro del entorno corporativo.

Desde una perspectiva técnica, el uso de una plantilla reduce significativamente el coste computacional asociado a la utilización de modelos de lenguaje. Una parte sustancial del

contenido de un informe de seguimiento es estructural y repetitiva, incluyendo descripciones de pruebas, explicaciones de escalas y encabezados estandarizados. Al mantener estos elementos predefinidos en el documento base, se evita su envío innecesario al modelo de Inteligencia Artificial, disminuyendo el número de tokens procesados y, en consecuencia, el coste económico por generación.

Asimismo, esta estrategia mejora el control semántico del sistema. La Inteligencia Artificial interviene únicamente en bloques concretos destinados a la interpretación evolutiva o síntesis de información, sin modificar la arquitectura global del documento. Finalmente, desde el punto de vista operativo, la utilización de plantillas permite adaptar el formato del informe ante posibles cambios organizativos o normativos sin necesidad de alterar la lógica del programa, lo que aporta flexibilidad y sostenibilidad al sistema.

### **8.3 SISTEMA DE ETIQUETAS DINÁMICAS**

La plantilla incorpora un sistema de marcadores identificables mediante el patrón <<ETIQUETA>>, que actúan como posiciones reservadas para la inserción dinámica de contenido. Estas etiquetas permiten delimitar con precisión los puntos del documento susceptibles de sustitución automática, evitando interferencias con el texto ordinario del informe gracias al uso de delimitadores dobles angulares fácilmente reconocibles.

Algunos ejemplos de etiquetas son: <<ENFERMERIA\_IA\_INTEGRIDADCUTANEA>>, <<NOMBRE>>, <<NORTON\_RESULTADO>> o <<NORTON\_TABLA>>.

La nomenclatura empleada sigue una convención estructurada que facilita tanto la lectura como el mantenimiento del sistema. En el caso de las escalas clínicas, se utiliza el patrón NOMBREPRUEBA\_TIPO, donde el sufijo identifica la naturaleza del contenido a insertar, como \_RESULTADO, \_TABLA o \_GRAFICO. Esta estructura permite distinguir de forma inequívoca si la etiqueta debe ser sustituida por un valor puntual, una tabla histórica o una representación visual. Para los bloques generados mediante Inteligencia Artificial se emplea una convención diferente basada en el departamento o área asistencial, seguida del

identificador IA y del ámbito concreto de análisis, como por ejemplo ENFERMERIA\_IA\_INTEGRIDAD CUTÁNEA o FISIO\_IA\_CAIDAS. Esta diferenciación explícita permite separar conceptualmente los contenidos derivados de datos estructurados de aquellos que requieren interpretación o síntesis.

Cada etiqueta está asociada internamente a un valor concreto que puede corresponder a un dato puntual, una estructura tabular, un gráfico generado previamente o un bloque interpretativo producido mediante Inteligencia Artificial. Durante la ejecución del sistema, el documento es recorrido de manera programática, identificando la presencia de estas marcas tanto en párrafos como en celdas de tablas. Una vez detectadas, se procede a su reemplazo por el contenido correspondiente.

Este mecanismo de etiquetado proporciona un alto grado de flexibilidad y escalabilidad. La incorporación de nuevos apartados o variables no requiere modificar la estructura general del código, sino únicamente añadir nuevas asociaciones entre etiquetas y contenidos. De este modo, la evolución futura del sistema puede realizarse de forma modular y controlada.

## **8.4 SUSTITUCIÓN DE CONTENIDO Y GESTIÓN DE EXCEPCIONES**

El sistema implementa distintos mecanismos de sustitución en función de la naturaleza del contenido asociado a cada etiqueta. Esta diferenciación permite adaptar el proceso de inserción al tipo de información que debe incorporarse al documento, garantizando coherencia formal y exactitud en el resultado final. Actualmente se permite realizar una sustitución tanto por texto simple como por tablas dinámicas, gráficas o bloques generados mediante IA.

### **8.4.1 SUSTITUCIÓN POR TEXTO SIMPLE**

Este tipo de sustitución se aplica a etiquetas asociadas a valores unitarios como nombres, fechas, resultados numéricos o pequeñas observaciones. El procedimiento consiste en recorrer los párrafos del documento Word, localizar la etiqueta correspondiente y reemplazarla por el valor calculado previamente.

```
def reemplazar_etiqueta_simple(doc, etiqueta, valor, formato=None):

    if formato is None:
        formato = 1

    colores = {
        1: RGBColor(0x93, 0x99, 0xC4), # Azul claro
        2: RGBColor(0x00, 0x00, 0x00), # Negro
        3: RGBColor(0xEA, 0xAA, 0x00), # Amarillo
        4: RGBColor(0xFF, 0x00, 0x00) # Rojo
    }

    color = colores.get(formato, colores[1])

    for p in doc.paragraphs:
        # 1) Unificar todo el texto del párrafo
        texto_original = "".join(run.text for run in p.runs)
        if etiqueta not in texto_original:
            continue
        texto_nuevo = texto_original.replace(etiqueta, str(valor))

        for run in list(p.runs):
            p._element.remove(run._element)

        nuevo_run = p.add_run(texto_nuevo)
        nuevo_run.font.color.rgb = color
        run.font.size = Pt(11)
```

Este mecanismo se aplica, por ejemplo, a etiquetas como: <<NOMBRE\_RESIDENTE>>, <<NORTON\_RESULTADO>> o <<EVA\_FECHA>>. A continuación, se muestra un ejemplo.

|                       |                          |
|-----------------------|--------------------------|
| - Fecha de la prueba: | <<NORTON_FECHA>>         |
| - Resultado:          | <<NORTON_RESULTADO>>     |
| - Valoración:         | <<NORTON_VALORACION>>    |
| - Observaciones:      | <<NORTON_OBSERVACIONES>> |

|                       |            |
|-----------------------|------------|
| - Fecha de la prueba: | 02-10-2025 |
| - Resultado:          | 12         |
| - Valoración:         | Alto       |
| - Observaciones:      | Sin obs.   |

*Figura 9 - Ejemplo de sustitución por texto simple*

## 8.4.2 SUSTITUCIÓN POR TABLAS DINÁMICAS

Cuando la etiqueta hace referencia a un histórico de datos, como la evolución de una escala clínica o un registro de incidencias, el sistema genera dinámicamente una tabla a partir de una estructura tipo DataFrame. Posteriormente, dicha tabla se inserta en el documento en la posición correspondiente.

```
def reemplazar_etiqueta_con_tabla(doc, etiqueta, dfs, formato=None,
    ampliar_campo=False):

    # 1) Normalizar dfs a lista
    if isinstance(dfs, dict):
        dfs = list(dfs.values())
    elif isinstance(dfs, pd.DataFrame):
        dfs = [dfs]
    elif not isinstance(dfs, (list, tuple)):
        raise TypeError("`dfs` debe ser un DataFrame, una lista/tupla de
DataFrames, o un dict de DataFrames")

    if formato is None:
        formato = 1

    colores = {
        1: RGBColor(0x93, 0x99, 0xC4), # Azul claro
        2: RGBColor(0x00, 0x00, 0x00), # Negro
        3: RGBColor(0xEA, 0xAA, 0x00), # Amarillo
        4: RGBColor(0xFF, 0x00, 0x00)  # Rojo
    }

    color = colores.get(formato, colores[1])

    # Cálculo de ancho total deseado
    section = doc.sections[0]
    usable_width = section.page_width - section.left_margin -
section.right_margin * 2
    #extra_width = Inches(1)
    total_width = usable_width# + extra_width

    for p in doc.paragraphs:
        if etiqueta in p.text:
            last_elt = p._element # inicio de inserción

            for df in dfs:
                # >>> Evitar dfs sin columnas
                if df.shape[1] == 0:
                    continue
                num_cols = len(df.columns)
                # --- Crear tabla ---
                tabla = doc.add_table(rows=1, cols=num_cols, style='Table Grid')
                tabla.alignment = WD_TABLE_ALIGNMENT.CENTER
```

```
tabla.autofit = False
tabla.width = total_width

# --- Calcular anchos por columna ---
if ampliar_campo:
    if ampliar_campo in df.columns:
        idx_ampliar = list(df.columns).index(ampliar_campo)
        ancho_grande = Inches(3)
        resto = total_width - ancho_grande
        if num_cols > 1:
            ancho_otras = resto / (num_cols - 1)
        else:
            ancho_otras = resto
        # construir lista de anchos
        anchos = [
            ancho_grande if j == idx_ampliar else ancho_otras
            for j in range(num_cols)
        ]
    else:
        print(f"Columna '{ampliar_campo}' no encontrada.")
        anchos = [total_width / num_cols] * num_cols
else:
    anchos = [total_width / num_cols] * num_cols

# aplicar anchos
for j, col in enumerate(tabla.columns):
    for cell in col.cells:
        cell.width = anchos[j]

# --- Cabecera ---
for j, col_name in enumerate(df.columns):
    cell = tabla.cell(0, j)
    cell.text = str(col_name)
    cell.vertical_alignment = WD_ALIGN_VERTICAL.CENTER
    for para in cell.paragraphs:
        para.alignment = WD_ALIGN_PARAGRAPH.CENTER
        for run in para.runs:
            run.font.color.rgb = color
            run.font.size = Pt(11)

# --- Filas de datos ---
for _, fila in df.iterrows():
    celdas = tabla.add_row().cells
    for j, valor in enumerate(fila):
        cell = celdas[j]
        cell.text = str(valor)
        cell.vertical_alignment = WD_ALIGN_VERTICAL.CENTER
        for para in cell.paragraphs:
            para.alignment = WD_ALIGN_PARAGRAPH.CENTER
            for run in para.runs:
                run.font.color.rgb = color
                run.font.size = Pt(11)
```

```
# insertar tabla tras el último elemento
last_elt.addnext(tabla._tbl)
last_elt = tabla._tbl

# párrafo en blanco como separación
blank = XmlElement('w:p')
last_elt.addnext(blank)
last_elt = blank

# limpiar marcador
p.text = p.text.replace(etiqueta, "")
return
```

Este procedimiento se aplica a etiquetas como: <<NORTON\_TABLA>>, <<DOWNTON\_TABLA>> o <<ENFERMERIA\_ULCERAS>>. A continuación se muestra un ejemplo.

Esta es la evolución del residente a lo largo de las <<EVA\_NUM\_PRUEBAS>> pruebas que ha realizado en el centro.

<<EVA\_TABLA>>

| Fecha      | Resultado | Observaciones                                                         |
|------------|-----------|-----------------------------------------------------------------------|
| 19-04-2024 | 0         | NO REFIERE DOLOR EN EL MOMENTO DE LA VALORACION                       |
| 09-10-2024 | 50        | REFIERE DOLOR A LA MOVILIZACION DE MID                                |
| 11-10-2024 | 50        | REFIERE DOLOR A LA MOVILIZACION DE MID                                |
| 22-01-2025 | 0         | No refiere dolor en este momento tras la operación de cadera derecha. |
| 10-03-2025 | 0         | No refiere dolor en este momento tras la operación de cadera derecha. |
| 03-04-2025 | 0         | No refiere dolor en este momento tras la operación de cadera derecha. |
| 27-06-2025 | 0         | No refiere dolor en este momento.                                     |
| 02-10-2025 | 0         | No refiere dolor en este momento.                                     |

Figura 10 - Ejemplo de sustitución por tabla dinámica

### 8.4.3 SUSTITUCIÓN POR GRÁFICAS

En los casos en que se requiere una representación visual de la evolución de una escala, el sistema genera previamente una gráfica mediante bibliotecas como Matplotlib y la guarda temporalmente como imagen. Posteriormente, dicha imagen se inserta en el documento en la posición de la etiqueta correspondiente.

```
def generar_grafico_escalas(df, grafico_path="grafico_resultado.png",
    columnas_extra=None):
    if columnas_extra is None:
        columnas_extra = []

    plt.figure(figsize=(6, 3))

    # Trazar columnas extra
    for col in columnas_extra:
        if col in df.columns:
            print(f"Columna {col} encontrada en escala, añadiendo al gráfico.")
            plt.plot(df["Fecha"], df[col].dropna(), marker='o', label=col)
        else:
            print(f"Columna {col} no encontrada en escala, omitiendo.")
    # Trazar resultado principal
    plt.plot(df["Fecha"], df["Resultado"], marker='o', label="Resultado")

    plt.title("Evolución del Resultado", y = 1.23)
    plt.xlabel("Fecha")
    plt.ylabel("Valor")
    plt.grid(True)
    plt.gca().xaxis.set_major_formatter(mdates.DateFormatter('%Y-%m-%d'))
    plt.gca().xaxis.set_major_locator(mdates.AutoDateLocator())
    plt.xticks(rotation=45, ha='right')
    plt.legend()
    # 1) Leyenda horizontal encima del gráfico
    plt.legend(
        loc='upper center',
        bbox_to_anchor=(0.5, 1.3),
        ncol=len(columnas_extra) + 1, # tantas columnas como series en la
    leyenda
        frameon=False
    )

    plt.tight_layout()
    plt.savefig(grafico_path)
    plt.close()
    return grafico_path
```

```
def reemplazar_etiqueta_con_imagen(doc, etiqueta, imagen_path):
```

```
for p in doc.paragraphs:
    if etiqueta in p.text:
        p.text = p.text.replace(etiqueta, "")
        p.add_run().add_picture(imagen_path, width=Inches(5.5))
    return
```

Este mecanismo se utiliza en etiquetas como: <<NORTON\_GRAFICO>>, <<EVA\_GRAFICO>> o <<TINETTI\_GRAFICO>>. A continuación se muestra un ejemplo.

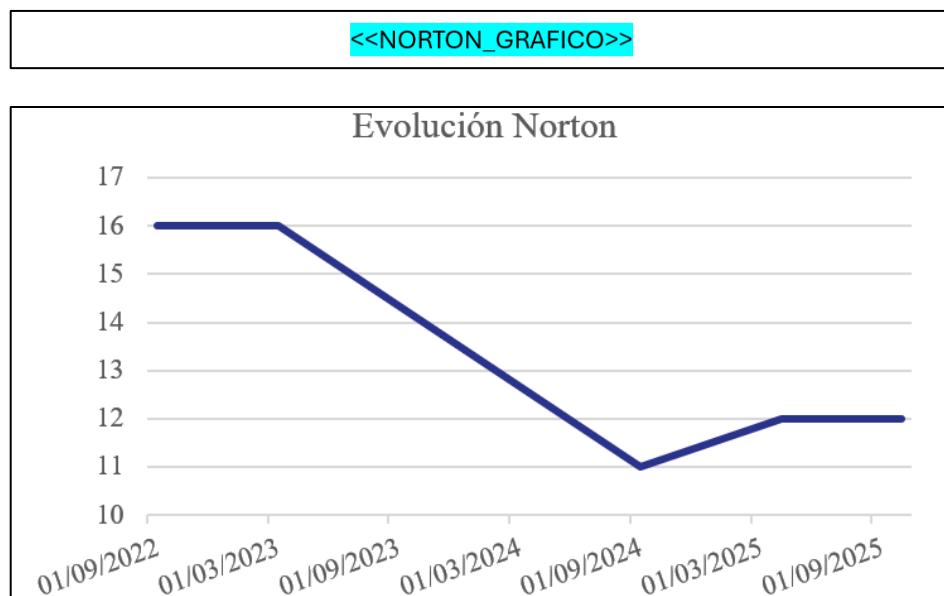


Figura 11 - Ejemplo de sustitución por gráfica

#### 8.4.4 SUSTITUCIÓN POR BLOQUES GENERADOS MEDIANTE IA

Los apartados interpretativos del informe, como la evolución funcional o cognitiva del residente, se generan previamente mediante el modelo de lenguaje bajo restricciones definidas en el Capítulo 9. Una vez obtenido el texto generado, este se inserta en la posición correspondiente del documento.

```
def generarIA_UnaSolaEscala(nombre_prueba, area, resultados_escalas, api_key,
nombre_residente = "No proporcionado", modelo="o3"):
    df = resultados_escalas[nombre_prueba]["datos"]
    descripcion = resultados_escalas[nombre_prueba]["descripcion"]
    if not df.empty:
```

```
print(f"\nGenerando resumen de '{area}' para la prueba  
'{nombre_prueba}'...")  
resumen = generar_resumen_observaciones_escalas(df, api_key,  
nombre_prueba, descripcion, area, nombre_residente, modelo)  
else:  
    resumen = f"Sin datos de {nombre_prueba} para generar resumen de {area}"  
return resumen
```

Etiquetas típicas de este tipo son: <<ENFERMERIA\_IA\_INTEGRIDADCUTANEA>>, <<FISIO\_IA\_ESCALASCAIDAS>> o <<PSICO\_IA\_VALORACIONCOGNITIVA>>. A continuación se muestra un ejemplo.

**Resumen de integridad cutánea con IA**

El residente presenta tres úlceras registradas en la residencia: una en el codo derecho, otra en la pierna izquierda y una en el brazo izquierdo.

En los últimos meses, las dos primeras han cicatrizado y ya no están activas.

La úlcera del brazo izquierdo permanece abierta y sigue en tratamiento.

El cuidado ha incluido betadine y silicona, con apósitos cada 24 horas para la zona afectada.

La escala Norton mostró varias valoraciones: dos resultados de 16 (riesgo mínimo) y otros de 11, 12, 12 y 14 (riesgo alto a medio), indicando fluctuación del riesgo.

En conjunto, la evolución es favorable para las heridas ya cerradas; la del brazo izquierdo requiere vigilancia continua para evitar complicaciones y favorecer su cierre.

*Figura 12 - Ejemplo de texto generado por IA*

#### 8.4.5 GESTIÓN DE EXCEPCIONES DOCUMENTALES

Para evitar errores silenciosos o la permanencia accidental de etiquetas sin reemplazar, se ha implementado un mecanismo de control que verifica la disponibilidad de datos antes de cada sustitución. En caso de no existir información asociada a una etiqueta determinada, el sistema inserta un texto explícito como: “No hay datos disponibles”

Este texto se inserta con formato diferenciado (por ejemplo, color rojo), permitiendo al profesional identificar visualmente la ausencia de información. De esta manera se evita la generación de contenido inventado, refuerza la trazabilidad del sistema y mantiene la transparencia del informe final.

```
if df.empty:
    reemplazar_etiqueta_simple(
        doc,
        etiqueta(nombre_prueba, "TABLA"),
        "No hay datos disponibles para esta escala.",
        formato=4 # Texto en rojo para advertencia
    )
else:
    reemplazar_etiqueta_con_tabla(
        doc,
        etiqueta(nombre_prueba, "TABLA"),
        df,
        formato=2
    )
```

```
if not df.empty:
    resumen = generar_resumen_observaciones_escalas(...)
else:
    resumen = "Sin datos suficientes para generar resumen interpretativo."
```

## **Capítulo 9. MODELO CONTROLADO DE INTELIGENCIA ARTIFICIAL**

En este capítulo se describe el diseño e implementación del modelo de Inteligencia Artificial utilizado para la generación asistida de informes de seguimiento del Plan de Atención Individualizada (PAI). El enfoque adoptado se basa en un uso controlado de modelos de lenguaje, cuyo objetivo no es la generación libre de contenido, sino la síntesis supervisada de información clínica previamente validada, garantizando la fiabilidad, coherencia y seguridad de los textos generados.

Dado el carácter sensible de la información tratada y su impacto potencial en la toma de decisiones asistenciales, el modelo ha sido diseñado para actuar como una herramienta de apoyo a la labor profesional, evitando en todo momento la generación de información no respaldada por los datos disponibles.

### **9.1 PRINCIPIOS DEL USO CONTROLADO DE LA IA**

Tal y como se ha expuesto en capítulos anteriores, uno de los principales riesgos asociados al uso de modelos de lenguaje de gran tamaño en entornos clínicos es la generación de información incorrecta o no verificable, fenómeno conocido como alucinación [21]. Para mitigar este riesgo, el sistema desarrollado restringe de forma estricta el contexto de actuación del modelo, limitando su entrada exclusivamente a datos reales procedentes de los sistemas corporativos.

El modelo no tiene acceso a información externa ni capacidad para inferir datos no proporcionados explícitamente. Su función se limita a interpretar y sintetizar la información disponible, actuando como un copiloto de redacción que asiste al profesional en la elaboración del informe, pero no sustituye su criterio clínico.

## **9.2 DISEÑO Y ESTRUCTURACIÓN DE PROMPTS**

El diseño de los prompts constituye un elemento clave del sistema, ya que determina tanto la calidad del contenido generado como el grado de control sobre el comportamiento del modelo. Los prompts han sido diseñados de forma exhaustiva, incorporando instrucciones explícitas sobre:

- El rol que debe asumir el modelo (profesional sociosanitario).
- La fuente exclusiva de la información (datos proporcionados).
- La prohibición explícita de inventar información.
- El formato y extensión de la respuesta.
- El lenguaje y tono del texto generado.
- El público destinatario del informe (residente y familia).

Un aspecto especialmente relevante es el uso deliberado del término “residente” en lugar de “paciente”, siguiendo las directrices internas de la organización y alineándose con un enfoque de atención centrada en la persona. Este matiz terminológico se incorpora de forma explícita en los prompts para garantizar la coherencia del lenguaje utilizado en los informes generados.

Asimismo, los prompts incluyen instrucciones específicas para la interpretación clínica, priorizando la descripción de cambios funcionales, cognitivos, emocionales o de salud frente a la simple enumeración de valores numéricos.

## **9.3 EJEMPLOS DE PROMPTS UTILIZADOS**

Cada uno de los prompts recibe con entrada un conjunto completo de datos estructurados y no estructurados asociados a una escala clínica completa. Estos datos incluyen la identificación del residente, descripción de la prueba o escala aplicada, resultados históricos de la escala, observaciones clínicas registradas por los profesionales o el área asistencial a la que pertenece la valoración.

Este enfoque permite al modelo disponer de una visión longitudinal y contextualizada de la evolución del residente, elemento fundamental para generar un resumen coherente y clínicamente relevante.

A continuación, se muestran dos ejemplos representativos de los prompts diseñados, correspondientes a la generación de resúmenes evolutivos a partir de escalas clínicas. En el segundo prompt, en el caso de las escalas EVA, el prompt contempla explícitamente situaciones en las que la prueba no puede ser aplicada, dando prioridad a escalas alternativas como PAINAD cuando esta información aparece reflejada en las observaciones clínicas. Este tipo de lógica incorporada en el diseño de los prompts refuerza el carácter controlado y clínicamente coherente del sistema.

Estas son las escalas clínicas pasadas a un residente (nunca "paciente") a lo largo del tiempo en Valdeluz Mayores, para cada vez que se pasa la escala {nombre\_prueba}, consistente en lo siguiente: {descripcion\_prueba}.

Desde la perspectiva de un profesional sociosanitario haz un resumen, sin inventarte ningún dato, de la evolución del residente teniendo en cuenta los resultados obtenidos, su evolución y las observaciones del profesional.

Redactado, con tamaño máximo de 8 líneas. Interpreta los resultados obtenidos y su evolución, no describas variaciones entre resultados numéricos que resultan evidentes, sino variaciones en el estado de salud, estado cognitivo, estado funcional o estado emocional del residente. No incluyas datos concretos de fechas, en todo caso "desde la última evaluación", "en los últimos tres meses", etc. Escribe en formato lista, por puntos, pero usando simplemente dobles guiones como identadores sin espacio después.

Usa lenguaje sencillo, teniendo en cuenta que es un informe sobre el estado de salud, cognitivo, funcional y emocional destinado al residente y su familia. No es un informe interno.

El nombre del residente es {nombre\_residente}, en caso de que sea distinto de "No proporcionado" utilízalo en tu respuesta de manera natural, sin apellido. Si es "No proporcionado", refiérete al residente siempre como "el residente". Si se proporciona el nombre de una mujer, usa el género femenino al escribir.

El nombre del área enfocada es {area\_enfocada}, tenlo en cuenta sin necesariamente referenciarlo directamente.

Refiérete al residente en tercera persona.

Observaciones:

{texto\_observaciones}

Resultados:

{', '.join(map(str, resultados))}

Estas son las escalas clínicas pasadas a un residente (nunca "paciente") a lo largo del tiempo en Valdeluz Mayores, para cada vez que se pasa la escala EVA, consistente en lo siguiente: "EVA es el nombre que recibe una escala horizontal utilizada para medir el dolor en el ámbito de la sanidad. La escala EVA Permite establecer una valoración del dolor que oscila entre 0 y 10, siendo el 0 la ausencia completa de dolor y el 10 un dolor insoportable."

Desde la perspectiva de un profesional sociosanitario haz un resumen, sin inventarte ningún dato, de la evolución del residente teniendo en cuenta los resultados obtenidos, su evolución y las observaciones del profesional.

Redactado, con tamaño máximo de 8 líneas. Interpreta los resultados obtenidos y su evolución, no describas variaciones entre resultados numéricos que resultan evidentes, sino variaciones en el estado de salud, estado cognitivo, estado funcional o estado emocional del residente. No incluyas datos concretos de fechas, en todo caso "desde la última evaluación", "en los últimos tres meses", etc. Escribe en formato lista, por puntos, pero usando simplemente dobles guiones como identadores sin espacio después.

Usa lenguaje sencillo, sin tecnicismos médicos, teniendo en cuenta que es un informe sobre el estado de salud, cognitivo, funcional y emocional destinado al residente y su familia. No es un informe interno.

El nombre del residente es `{nombre_residente}`, en caso de que sea distinto de "No proporcionado" utilízalo en tu respuesta de manera natural, sin apellido. Si es "No proporcionado", refiérrete al residente siempre como "el residente". Si se proporciona el nombre de una mujer, usa el género femenino al escribir. Refiérrete al residente en tercera persona.

El nombre del área enfocada es "Valoración del dolor", tenlo en cuenta sin necesariamente referenciarlo directamente.

Observaciones:

`{texto_observaciones}`

Resultados:

`{', '.join(map(str, resultados))}`

Ten en cuenta que la escala EVA no siempre se puede pasar. En ese caso se avisaría en las Observaciones, y en ocasiones aparece en ese mismo campo el resultado de la escala PAINAD, mucho más precisa. Siempre tiene prioridad sobre la EVA. Su descripción es la siguiente:

La escala Pain Assessment in Advanced Dementia (PAINAD) valora el dolor en personas con trastornos de comunicación y con demencia. Es una escala de fácil aplicación que consta de 5 elementos: respiración, vocalización negativa, expresión facial, lenguaje corporal y capacidad de alivio. Cada ítem puede obtener una puntuación máxima de 2. Las puntuaciones totales pueden ser de 0 (sin dolor) a 10 (máximo dolor), equivalente a la escala EVA.

## **Capítulo 10. ANÁLISIS DE RESULTADOS**

El presente capítulo tiene como objetivo evaluar los resultados obtenidos tras el desarrollo e implementación del sistema de generación asistida de informes del Plan de Atención Individualizada (PAI). El análisis se aborda desde una perspectiva funcional, operativa y cualitativa, permitiendo valorar tanto el rendimiento técnico del sistema como su impacto en la práctica profesional.

A lo largo de este capítulo se examina el grado de cumplimiento de los objetivos planteados, la eficiencia del sistema en términos de tiempo y recursos, la percepción por parte de los profesionales usuarios y las principales limitaciones identificadas durante su utilización.

### ***10.1 VALIDACIÓN FUNCIONAL DEL SISTEMA***

Los resultados obtenidos confirman que el sistema es capaz de recorrer de forma íntegra todas las etapas del proceso: desde la obtención de datos hasta la generación final del documento en formato Word. En todos los casos analizados, el sistema ha logrado integrar correctamente los distintos tipos de contenido, incluyendo datos simples, tablas históricas, representaciones gráficas y bloques interpretativos generados mediante Inteligencia Artificial.

Un aspecto especialmente relevante ha sido la correcta sustitución de las etiquetas definidas en la plantilla. No se han detectado incidencias relacionadas con la permanencia de marcadores sin reemplazar, lo que confirma la robustez del mecanismo de generación documental. Asimismo, el sistema mantiene la coherencia de formato en todos los elementos insertados, respetando la estructura original del documento.

En aquellos casos en los que no existía información disponible, el sistema ha respondido adecuadamente mediante la inserción de mensajes explícitos que informan de la ausencia de

datos. Este comportamiento resulta fundamental para garantizar la transparencia del informe y evitar la generación de contenido no respaldado por información real.

En conjunto, la validación funcional demuestra que el sistema cumple de forma satisfactoria con los requisitos definidos, ofreciendo un comportamiento estable y predecible en distintos escenarios.

## ***10.2 EVALUACIÓN DE EFICIENCIA Y TIEMPO***

Uno de los principales objetivos del sistema es optimizar el tiempo dedicado a la elaboración de informes PAI, una tarea que tradicionalmente implica un proceso manual intensivo basado en la revisión de datos, su interpretación y la redacción del documento final.

En el modelo tradicional, la elaboración de un informe requiere una inversión significativa de tiempo por parte del profesional, especialmente en casos donde existe un histórico amplio de valoraciones. Este proceso incluye la consulta de múltiples fuentes de información, la síntesis de resultados y la construcción manual del documento, lo que puede dar lugar a variabilidad en el formato y en la calidad del contenido.

Con la implantación del sistema desarrollado, este flujo se automatiza en gran medida. La generación del informe completo se realiza en un tiempo significativamente reducido, limitándose la intervención del profesional a la revisión final del contenido. Esta reducción del tiempo operativo no solo mejora la eficiencia, sino que también permite a los profesionales dedicar más tiempo a tareas asistenciales.

Desde el punto de vista técnico, la eficiencia del sistema se ve reforzada por el uso de un modelo híbrido que limita el uso de la Inteligencia Artificial a los bloques estrictamente necesarios. Al evitar la generación completa del documento mediante IA, se reduce el consumo de recursos computacionales y el coste asociado al uso del modelo de lenguaje.

En términos generales, el sistema consigue una mejora sustancial en la productividad, manteniendo al mismo tiempo un alto nivel de calidad y coherencia en los informes generados.

### ***10.3 VALORACIÓN POR PARTE DE LOS PROFESIONALES***

La evaluación cualitativa del sistema se ha realizado a partir de la percepción de los profesionales sociosanitarios que han interactuado con los informes generados. Esta valoración resulta clave, ya que el éxito del sistema no depende únicamente de su funcionamiento técnico, sino también de su aceptación en el entorno asistencial.

Los profesionales han destacado especialmente la reducción del tiempo necesario para la elaboración de informes, así como la mejora en la homogeneidad documental. La existencia de una estructura fija y coherente facilita la lectura, la comparación entre informes y su integración en los sistemas corporativos.

Otro aspecto valorado positivamente ha sido la claridad del lenguaje utilizado en los textos generados mediante Inteligencia Artificial. El hecho de que los informes estén orientados a residentes y familiares, y no exclusivamente a profesionales, ha contribuido a mejorar la comprensión del contenido.

Asimismo, el diseño del sistema, en el que la Inteligencia Artificial actúa como una herramienta de apoyo y no como un sustituto del profesional, ha favorecido la generación de confianza. La posibilidad de identificar visualmente los distintos tipos de contenido refuerza la percepción de control sobre el informe.

En términos generales, la valoración por parte de los profesionales puede considerarse favorable, destacando su utilidad práctica y su integración natural en el flujo de trabajo habitual.

## ***10.4 LIMITACIONES DETECTADAS***

A pesar de los resultados positivos obtenidos, el sistema presenta una serie de limitaciones que deben ser consideradas para una correcta interpretación de su alcance.

En primer lugar, el sistema depende directamente de la calidad y disponibilidad de los datos. En aquellos casos en los que la información es incompleta o escasa, la capacidad del sistema para generar contenido relevante se ve limitada. Esta dependencia pone de manifiesto la importancia de mantener registros clínicos completos y actualizados.

Por otro lado, aunque el uso de la Inteligencia Artificial está restringido y controlado, la calidad del contenido generado depende en gran medida del diseño de los prompts y de la riqueza de los datos de entrada. En situaciones con poca información contextual, los textos generados pueden resultar menos precisos o útiles desde el punto de vista clínico.

Otra limitación identificada es la relativa homogeneidad en la redacción de los informes. Si bien esta característica contribuye a la estandarización, también puede generar cierta repetitividad en documentos sucesivos.

Desde el punto de vista técnico, el sistema ha sido diseñado para un entorno específico, lo que implica que su adaptación a otros contextos requeriría ajustes en la estructura de datos, en las plantillas y en los modelos utilizados.

Finalmente, es importante destacar que el sistema no sustituye en ningún caso el criterio clínico del profesional. La validación final del informe sigue siendo un paso imprescindible, especialmente en situaciones complejas o con información ambigua.

## **Capítulo 11. CONCLUSIONES Y TRABAJOS**

### **FUTUROS**

En este capítulo se presentan las conclusiones principales del Trabajo de Fin de Máster, sintetizando los objetivos alcanzados y las aportaciones realizadas, así como las líneas de evolución y mejora que se abren a partir del sistema desarrollado.

#### ***11.1 CONCLUSIONES***

El presente Trabajo de Fin de Máster ha abordado una problemática real y ampliamente extendida en el ámbito sociosanitario: la elevada carga administrativa asociada a la elaboración manual de informes de seguimiento del Plan de Atención Individualizada (PAI), a pesar de la disponibilidad de grandes volúmenes de datos registrados en los sistemas de información corporativos.

A lo largo del proyecto se ha diseñado e implementado un sistema de generación asistida de informes que integra datos procedentes del ERP ResiPlus y del datalake corporativo, combinando arquitecturas modernas de datos con técnicas de Inteligencia Artificial aplicadas de forma controlada. El sistema desarrollado permite transformar información clínica, funcional y asistencial previamente registrada en borradores estructurados de informes PAI, reduciendo significativamente el esfuerzo manual necesario para su redacción.

Los objetivos planteados al inicio del proyecto han sido alcanzados de forma satisfactoria. En particular, se ha logrado analizar y estructurar los datos relevantes para el seguimiento del residente, diseñar un proceso de extracción y tratamiento de la información, definir y validar prompts específicos para la síntesis clínica, y desarrollar un mecanismo de generación de informes basado en plantillas editables que facilita la validación por parte de

los profesionales. Asimismo, se han incorporado mecanismos de control orientados a minimizar errores o inferencias no fundamentadas, garantizando que la Inteligencia Artificial actúe como herramienta de apoyo y no como sustituto del criterio clínico.

Desde un punto de vista organizativo y asistencial, el sistema aporta un valor significativo al liberar tiempo que anteriormente se dedicaba a tareas repetitivas de síntesis documental. Este tiempo no se traduce en un ahorro salarial directo, sino en una reasignación hacia actividades de mayor valor asistencial, como la atención directa al residente, el seguimiento clínico continuado o la coordinación interdisciplinar. En este sentido, la herramienta contribuye a mejorar la eficiencia de los procesos y la calidad global de la atención prestada, alineándose con los principios de atención centrada en la persona.

Por último, cabe destacar que el coste de implantación y mantenimiento del sistema resulta reducido y sostenible en el tiempo, especialmente en comparación con el valor organizativo y asistencial generado. Esto refuerza la viabilidad de la solución y su potencial de escalado a otros centros residenciales.

## ***11.2 TRABAJOS FUTUROS***

A partir del trabajo realizado, se identifican diversas líneas de mejora y ampliación que podrían abordarse en el futuro. En primer lugar, sería de interés extender el sistema a otros tipos de informes clínicos o asistenciales, más allá del PAI, aprovechando la flexibilidad de la arquitectura y del modelo de generación de contenidos desarrollado.

Asimismo, podrían incorporarse mecanismos avanzados de personalización de los informes, adaptando el nivel de detalle y el lenguaje empleado en función del perfil del destinatario (familiares, profesionales internos o responsables del centro). En esta línea, la inclusión de métricas de calidad percibida por parte de los usuarios finales permitiría refinar progresivamente los prompts y los criterios de generación.

Otra línea relevante de trabajo futuro sería la integración del sistema en los flujos operativos habituales del centro, automatizando aún más la ejecución de los procesos y facilitando su

uso desde interfaces integradas en el propio sistema de gestión. Del mismo modo, la evaluación del sistema en un entorno de producción real, con un mayor volumen de residentes y centros, permitiría analizar su impacto a mayor escala y validar su comportamiento en escenarios más exigentes.

Finalmente, el uso de técnicas adicionales de Inteligencia Artificial, como modelos específicos entrenados sobre datos anonimizados del dominio sociosanitario, podría explorarse como complemento a los modelos generalistas, siempre manteniendo un enfoque ético, controlado y alineado con la normativa vigente.

## Capítulo 12. BIBLIOGRAFÍA

- [1] Python Software Foundation. (s.f.). *Python 3 documentation*. <https://docs.python.org/3/>.
- [2] GeeksforGeeks. (s.f.). SQL manual. <https://cdncontribute.geeksforgeeks.org/wp-content/uploads/SQL-Manual.pdf>.
- [3] Microsoft. (s.f.). Visual Studio Code documentation. <https://code.visualstudio.com/docs>.
- [4] Microsoft. (s.f.). Pylance. PyPI. <https://pypi.org/project/pylance/>.
- [5] Project Jupyter. (s.f.). Project Jupyter documentation. <https://docs.jupyter.org/en/latest/>.
- [6] Atlassian. (s.f.). Bitbucket guides. <https://bitbucket.org/product/guides>.
- [7] Chacon, S., & Straub, B. (s.f.). Pro Git. Git-scm. <https://git-scm.com/book/en/v2>.
- [8] Databricks. (s.f.). Databricks documentation. <https://docs.databricks.com/aws/en/>.
- [9] Aphelion Soluciones Informáticas. (2026). *ResiPlus: software de gestión para centros de mayores*. <https://addinformatica.com/software-gestion-centros-mayores/>.
- [10] Databricks. (s.f.). Data import with JDBC. <https://assets.docs.databricks.com/extras/notebooks/source/data-import/jdbc.html>.
- [11] JPyype Team. (s.f.). JPyype documentation. <https://jpyype.readthedocs.io/en/latest/>.
- [12] PyPI. (s.f.). JayDeBeApi. <https://pypi.org/project/JayDeBeApi/>.
- [13] IBM. (2024). *¿Qué son los modelos de lenguaje de gran tamaño (LLM)?* IBM Think. <https://www.ibm.com/es-es/think/topics/large-language-models>.
- [14] Fundación Caser. (2015). *La atención sociosanitaria en España: situación actual y perspectivas de futuro*. <https://www.fundacioncaser.org/sites/default/files/adjuntos/2201.pdf>.
- [15] INE. (2022, 13 de octubre). *Proyecciones de población 2022-2072* [Nota de prensa]. Instituto Nacional de Estadística. [https://www.ine.es/prensa/pp\\_2022\\_2072.pdf](https://www.ine.es/prensa/pp_2022_2072.pdf).
- [16] NHS England. (2022). *What is integrated care?* <https://www.england.nhs.uk/integratedcare/what-is-integrated-care/>.
- [17] Bundesministerium für Gesundheit. (2025). *Long-term care insurance: support for people who need long-term care*. gesund.bund.de. <https://gesund.bund.de/en/pflegeversicherung>.
- [18] OECD. (2023). *Health at a Glance 2023: OECD Indicators*. OECD Publishing. [https://www.oecd.org/en/publications/health-at-a-glance-2023\\_7a7afb35-en.html](https://www.oecd.org/en/publications/health-at-a-glance-2023_7a7afb35-en.html).

- [19] Gobierno de España. (2006). *Ley 39/2006, de 14 de diciembre, de Promoción de la Autonomía Personal y Atención a las Personas en Situación de Dependencia*. Boletín Oficial del Estado, núm. 299. <https://www.boe.es/buscar/act.php?id=BOE-A-2006-21990>.
- [20] Comunidad de Madrid. (2024). *Orden de 14 de noviembre de 2024 por la que se regula el Plan de Atención Individualizada en centros residenciales para personas mayores*. Boletín Oficial de la Comunidad de Madrid. [https://www.bocm.es/boletin/CM\\_Orden\\_BOCM/2024/11/16/BOCM-20241116-1.PDF](https://www.bocm.es/boletin/CM_Orden_BOCM/2024/11/16/BOCM-20241116-1.PDF).
- [21] Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1–38. <https://dl.acm.org/doi/10.1145/3571730>.
- [22] Unión Europea. (2024). *Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial)*. Diario Oficial de la Unión Europea. <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:02024R1689-20240712>.
- [23] Metodología Agile: la revolución en las formas de trabajo. BBVA. <https://www.bbva.com/es/innovacion/metodologia-agile-la-revolucion-las-formas-trabajo/>.
- [24] Gantt Chart Basics. Asana. <https://asana.com/es/resources/gantt-chart-basics>.
- [25] Naciones Unidas. (s.f.). Objetivos y metas de desarrollo sostenible. Objetivos de Desarrollo Sostenible. <https://www.un.org/sustainabledevelopment/es/sustainable-development-goals/>.
- [26] Naciones Unidas. (s.f.). Trabajo decente y crecimiento económico. Objetivos de Desarrollo Sostenible. <https://www.un.org/sustainabledevelopment/es/economic-growth/>.
- [27] Naciones Unidas. (s.f.). Industria, innovación e infraestructura. Objetivos de Desarrollo Sostenible. <https://www.un.org/sustainabledevelopment/es/infrastructure/>.
- [28] Naciones Unidas. (s.f.). Objetivo 10: Reducir la desigualdad en y entre los países. <https://www.un.org/sustainabledevelopment/es/inequality/>.
- [29] Naciones Unidas. (s.f.). Objetivo 16: Paz, justicia e instituciones sólidas. <https://www.un.org/sustainabledevelopment/es/peace-justice/>.
- [30] Pandas. (2026). Pandas documentation. <https://pandas.pydata.org/docs/>.
- [31] OpenAI. (2026). Python API reference. OpenAI API. <https://developers.openai.com/api/reference/python>.
- [32] Python-docx. (2026). Python-docx documentation. <https://python-docx.readthedocs.io/en/latest/>.

- [33] Matplotlib. (2026). Matplotlib: Visualization with Python. <https://matplotlib.org/stable/index.html>.
- [34] SourceForge. (s.f.). JayDeBeApi. PyPI. <https://pypi.org/project/JayDeBeApi/>.
- [35] The JPyte Team. (s.f.). JPyte documentation. <https://jpyte.readthedocs.io/en/latest/>.
- [36] Typing. (2026). Support for type hints. Python 3.12.4 documentation. <https://docs.python.org/3/library/typing.html>.
- [37] Dataclasses. (2026). Data classes. Python 3.12.4 documentation. <https://docs.python.org/3/library/dataclasses.html>.
- [38] OS. (2026). Miscellaneous operating system interfaces. Python 3.12.4 documentation. <https://docs.python.org/3/library/os.html>.
- [39] RE. (2026). Regular expression operations. Python 3.12.4 documentation. <https://docs.python.org/3/library/re.html>.
- [40] JSON. (2026). JSON encoder and decoder. Python 3.12.4 documentation. <https://docs.python.org/3/library/json.html>.
- [41] Warnings. (2026). Warning control. Python 3.12.4 documentation. <https://docs.python.org/3/library/warnings.html>.
- [42] Sanitas. (s.f.). Mini-examen cognoscitivo de Lobo (MEC). <https://www.sanitas.es/biblioteca-de-salud/tercera-edad/demencias-y-patologias/demencias-y-deterioro-cognitivo/escala-mini-lobo>.
- [43] Sanitas. (s.f.). Escala de Barthel: valoración de la independencia. <https://www.sanitas.es/biblioteca-de-salud/tercera-edad/demencias-y-patologias/demencias-y-deterioro-cognitivo/escala-barthel>.
- [44] Sanitas. (s.f.). Escala de valoración de la marcha y el equilibrio de Tinetti. <https://www.sanitas.es/biblioteca-de-salud/tercera-edad/rehabilitacion-y-movilidad/escala-tinetti>.
- [45] Sanitas. (s.f.). Escala de Norton para úlceras por presión. <https://www.sanitas.es/biblioteca-de-salud/tercera-edad/demencias-y-patologias/otras-patologias-y-enfermedades/escala-norton-ulceras>.
- [46] Sanitas. (s.f.). Escala de valoración de riesgo de caídas de Downton. <https://www.sanitas.es/biblioteca-de-salud/tercera-edad/rehabilitacion-y-movilidad/escala-downton>.
- [47] La Fisioterapia. (s.f.). Timed Up and Go (TUG). <https://lafisioterapia.net/timed-up-and-go-tug/>.

- [48] Nesplora. (s.f.). Escala GDS (Geriatric Depression Scale). <https://nesplora.com/blog/escala-gds/>.
- [49] Sanitas. (s.f.). Escala EVA: medición del dolor. <https://www.sanitas.es/biblioteca-de-salud/enfermedades-y-trastornos/pruebas-y-diagnosticos/escala-eva>.
- [50] Consejo General de la Psicología de España. (2021). Guía para la evaluación de la depresión geriátrica (GDS). [https://www.cop.es/uploads/PDF/2021/GDS\\_evaluacion\\_COP.pdf](https://www.cop.es/uploads/PDF/2021/GDS_evaluacion_COP.pdf).
- [51] University of Washington. (s.f.). NPI (Neuropsychiatric Inventory) - UDS3. <https://files.alz.washington.edu/documentation/uds3-tip-b5-sp.pdf>.
- [52] Nestlé Nutrition Institute. (2021). Mini Nutritional Assessment (MNA). <https://www.mna-elderly.com/sites/default/files/2021-10/MNA-spanish.pdf>.
- [53] Sanitas. (s.f.). Escala de valoración del dolor PAINAD. <https://www.sanitas.es/biblioteca-de-salud/tercera-edad/demencias-y-patologias/demencias-y-deterioro-cognitivo/escala-painad-que-es>.

## ANEXO I: ALINEACIÓN DEL PROYECTO CON LOS ODS

Los Objetivos de Desarrollo Sostenible (ODS) [25] constituyen un conjunto de 17 metas globales definidas por las Naciones Unidas en el marco de la Agenda 2030 (Tabla 7), orientadas a promover un desarrollo equilibrado desde el punto de vista social, económico y ambiental. Estos objetivos abordan desafíos fundamentales como la salud y el bienestar, la reducción de desigualdades, la innovación tecnológica y el fortalecimiento de las instituciones, estableciendo un marco común de actuación para gobiernos, organizaciones y entidades privadas. En el ámbito de los proyectos de ingeniería, la alineación con los ODS permite evaluar el impacto social y ético de las soluciones desarrolladas, asegurando que la aplicación de la tecnología contribuya de forma responsable a la mejora de la calidad de vida y al desarrollo sostenible.

*Tabla 7 - Listado de los ODS*

| <i>ODS</i>               | <i>Objetivo</i>                                                                               |
|--------------------------|-----------------------------------------------------------------------------------------------|
| ODS 1. Fin de la pobreza | Erradicar la pobreza en todas sus formas y en todo el mundo.                                  |
| ODS 2. Hambre cero       | Poner fin al hambre, lograr la seguridad alimentaria y promover una agricultura sostenible.   |
| ODS 3. Salud y bienestar | Garantizar una vida sana y promover el bienestar para todas las personas en todas las edades. |

| <i>ODS</i>                                     | <i>Objetivo</i>                                                                                                 |
|------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|
| ODS 4. Educación de calidad                    | Garantizar una educación inclusiva, equitativa y de calidad y promover oportunidades de aprendizaje permanente. |
| ODS 5. Igualdad de género                      | Lograr la igualdad entre los géneros y empoderar a todas las mujeres y niñas.                                   |
| ODS 6. Agua limpia y saneamiento               | Garantizar la disponibilidad y la gestión sostenible del agua y el saneamiento para todos.                      |
| ODS 7. Energía asequible y no contaminante     | Garantizar el acceso a una energía asequible, segura, sostenible y moderna.                                     |
| ODS 8. Trabajo decente y crecimiento económico | Promover el crecimiento económico sostenido, inclusivo y el empleo digno para todos.                            |
| ODS 9. Industria, innovación e infraestructura | Desarrollar infraestructuras resilientes y fomentar la innovación y la industrialización sostenible.            |
| ODS 10. Reducción de las desigualdades         | Reducir la desigualdad dentro de los países y entre ellos.                                                      |
| ODS 11. Ciudades y comunidades sostenibles     | Lograr que las ciudades y comunidades sean inclusivas, seguras, resilientes y sostenibles.                      |
| ODS 12. Producción y consumo responsables      | Garantizar modalidades de consumo y producción sostenibles.                                                     |
| ODS 13. Acción por el clima                    | Adoptar medidas urgentes para combatir el cambio climático y sus efectos.                                       |

| <i>ODS</i>                                    | <i>Objetivo</i>                                                                                         |
|-----------------------------------------------|---------------------------------------------------------------------------------------------------------|
| ODS 14. Vida submarina                        | Conservar y utilizar de forma sostenible los océanos, mares y recursos marinos.                         |
| ODS 15. Vida de ecosistemas terrestres        | Proteger y restaurar los ecosistemas terrestres y detener la pérdida de biodiversidad.                  |
| ODS 16. Paz, justicia e instituciones sólidas | Promover sociedades pacíficas, justas e inclusivas con instituciones eficaces.                          |
| ODS 17. Alianzas para lograr los objetivos    | Fortalecer los medios de implementación y revitalizar la alianza mundial para el desarrollo sostenible. |

El proyecto desarrollado se alinea con 4 de los ODS explicados anteriormente. Se alinea directamente con el ODS 3: Salud y bienestar [26], al contribuir a la mejora de la calidad asistencial en centros residenciales mediante la optimización del seguimiento clínico de las personas mayores. La automatización asistida de los informes de seguimiento (PAI) permite a los profesionales sanitarios reducir la carga administrativa asociada a la síntesis manual de información, dedicando más tiempo al análisis clínico y a la atención directa del residente. Asimismo, la integración estructurada de datos procedentes de diferentes áreas asistenciales favorece una visión global del estado de salud del residente, contribuyendo a una atención más personalizada, continua y segura.

También contribuye al ODS 9: Industria, innovación e infraestructura [27] mediante la aplicación de tecnologías avanzadas de análisis de datos e inteligencia artificial en el ámbito sociosanitario. La utilización de arquitecturas basadas en datalake, servicios en la nube e integración con sistemas corporativos como ResiPlus supone un ejemplo de innovación tecnológica orientada a la modernización de infraestructuras digitales. Además, el desarrollo de una solución escalable y modular fomenta la adopción de prácticas innovadoras que

pueden ser replicadas en otros centros o entornos asistenciales, impulsando la transformación digital del sector.

En cuanto al ODS 10: Reducción de las desigualdades [28] se refleja en el objetivo del proyecto de garantizar un seguimiento homogéneo y sistemático de todos los residentes, independientemente del centro o del profesional que realice la atención. La estandarización de los informes PAI, basada en datos objetivos y estructurados, contribuye a minimizar variaciones en la calidad del seguimiento clínico derivadas de factores organizativos o de carga asistencial. De este modo, la herramienta favorece una atención más equitativa y reduce posibles desigualdades en el acceso a una evaluación integral y continuada de la salud de las personas mayores.

Finalmente, también se vincula con el ODS 16: Paz, justicia e instituciones sólidas [29] a través del refuerzo de la transparencia, la trazabilidad y la fiabilidad de los procesos de gestión clínica. El uso controlado de la inteligencia artificial, combinado con mecanismos de validación humana y cumplimiento del RGPD, contribuye al desarrollo de sistemas institucionales más sólidos y responsables. Asimismo, la correcta gestión de la información clínica y la reducción del riesgo de errores en los informes de seguimiento fortalecen la confianza en las instituciones asistenciales y en los sistemas tecnológicos que las sustentan.

## ANEXO II: LIBRERÍAS UTILIZADAS

En la Tabla 8 se muestran todas las librerías que se han utilizado en este proyecto, incluyendo una breve descripción del uso que se le ha dado a cada una.

Tabla 8 - Librerías utilizadas

| <i>Librería</i>  | <i>Tipo</i>           | <i>Función</i>                                                                                                                                                                                                                                                |
|------------------|-----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Pandas [30]      | Manipulación de datos | Manipulación y análisis de datos mediante estructuras tipo DataFrame, utilizada para leer resultados de consultas SQL, procesar el Excel de códigos y transformar la información de cada sección del informe.                                                 |
| OpenAI [31]      | IA / API externa      | Interfaz para la interacción con modelos de lenguaje, utilizada para generar de forma automática los resúmenes y valoraciones cualitativas de las distintas secciones (enfermería, fisioterapia, psicología, terapia ocupacional, animación, trabajo social). |
| Python-docx [32] | Generación documental | Creación y modificación de documentos en formato Microsoft Word, incluyendo inserción de texto, tablas, imágenes y estilos a partir de la plantilla del PAI.                                                                                                  |
| Matplotlib [33]  | Visualización         | Generación de gráficos (evolución de escalas en el tiempo) que posteriormente se insertan como imágenes en el documento Word.                                                                                                                                 |

| <i>Librería</i>     | <i>Tipo</i>                    | <i>Función</i>                                                                                                                                |
|---------------------|--------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
| JayDeBeApi<br>[34]  | Conectividad de bases de datos | Establecimiento de la conexión JDBC con la base de datos SQL Server para la ejecución de consultas.                                           |
| JPyPe [35]          | Estándar de Python             | Gestión y manipulación de fechas y horas, permitiendo el tratamiento temporal de los datos del informe.                                       |
| Typing [36]         | Estándar de Python             | Definición de tipos estáticos en Python (List, Dict, Any, Tuple, Optional) para mejorar la legibilidad, mantenibilidad y robustez del código. |
| Dataclasses<br>[37] | Estándar de Python             | Definición simplificada de clases de datos utilizadas como estructuras internas del proyecto.                                                 |
| Os [38]             | Estándar de Python             | Gestión de rutas de archivos y directorios del sistema operativo.                                                                             |
| Re [39]             | Estándar de Python             | Uso de expresiones regulares para la búsqueda y limpieza de texto dentro de los documentos y datos procesados.                                |
| Json [40]           | Estándar de Python             | Serialización y deserialización de datos en formato JSON, empleado en la comunicación con la API de OpenAI.                                   |
| Warnings<br>[41]    | Estándar de Python             | Gestión y supresión de advertencias generadas durante la ejecución del código.                                                                |

## ANEXO III: ESCALAS CLÍNICAS

En este anexo se definen las distintas escalas utilizadas para el desarrollo de este proyecto, así como los valores y rangos de puntuación junto con el significado.

### ***III.1 ESCALA DE LOBO (MEC - MINI-EXAMEN COGNOSCITIVO)***

La escala de Lobo [42] mide el estado cognitivo global, evaluando la orientación (temporal y espacial), la capacidad de fijación, la atención, el cálculo, la memoria a corto plazo, el lenguaje y la capacidad viso-constructiva. Es la herramienta de cribado de demencias más utilizada en España.

Consiste en una entrevista estructurada compuesta por una serie de preguntas y pequeñas tareas cognitivas. Algunos ejemplos son indicar la fecha actual o el lugar en el que se encuentra, repetir y recordar una lista de palabras, realizar operaciones sencillas (por ejemplo, restar de siete en siete) o copiar un dibujo simple.

La administración de esta prueba es directa por parte del profesional y suele requerir entre 10 y 15 minutos.

La puntuación sigue estos rangos (sobre 35 puntos):

- 0 a 23 puntos: Sospecha de deterioro cognitivo.
- 24 a 27 puntos: Posible deterioro leve o zona de duda.
- 28 a 35 puntos: Funcionamiento cognitivo normal.

Las preguntas se pueden ver en las siguientes tablas:

*Tabla 9 - Preguntas Escala Lobo de Orientación temporal*

| <i>Pregunta</i>      | <i>Puntuación</i> |
|----------------------|-------------------|
| ¿En qué día estamos? | 1                 |
| ¿En qué fecha?       | 1                 |
| ¿En qué mes?         | 1                 |
| ¿En qué estación?    | 1                 |
| ¿En qué año?         | 1                 |

*Tabla 10 - Preguntas Escala Lobo de Orientación espacial*

| <i>Pregunta</i>                        | <i>Puntuación</i> |
|----------------------------------------|-------------------|
| ¿En qué hospital o lugar estamos?      | 1                 |
| ¿En qué piso o planta?                 | 1                 |
| ¿En qué pueblo o ciudad?               | 1                 |
| ¿En qué provincia, región o autonomía? | 1                 |
| ¿En qué país?                          | 1                 |

*Tabla 11 - Preguntas Escala Lobo de Fijación*

| <i>Pregunta</i>                                                                                                                               | <i>Puntuación</i> |
|-----------------------------------------------------------------------------------------------------------------------------------------------|-------------------|
| Repita estas 3 palabras: Peseta - Caballo - Manzana. En caso de reevaluaciones se pueden utilizar alternativas como Libro – Queso - Bicicleta | 3                 |

*Tabla 12 - Preguntas Escala Lobo de Concentración y cálculo*

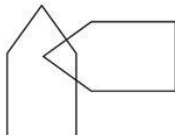
| <i>Pregunta</i>                                                   | <i>Puntuación</i> |
|-------------------------------------------------------------------|-------------------|
| Si tiene 30 ptas y le van quitando de 3 en 3, ¿cuántas le quedan? | 5                 |
| Repita 5-9-2 (hasta que los aprenda). Ahora hacia atrás.          | 3                 |

*Tabla 13 - Preguntas Escala Lobo de Memoria*

| <i>Pregunta</i>                                 | <i>Puntuación</i> |
|-------------------------------------------------|-------------------|
| ¿Recuerda las 3 palabras que le he dicho antes? | 3                 |

*Tabla 14 - Preguntas Escala Lobo de Lenguaje y construcción*

| <i>Pregunta</i>                                                                                       | <i>Puntuación</i> |
|-------------------------------------------------------------------------------------------------------|-------------------|
| Mostrarle un lápiz o un bolígrafo: ¿Qué es esto? Repetirlo con un reloj.                              | 2                 |
| Repita la frase ‘En un trigal había 5 perros’.                                                        | 1                 |
| Una manzana y una pera son frutas, ¿verdad? ¿Qué son un perro y un gato? ¿Qué son el verde y el rojo? | 2                 |
| Coja este papel con la mano derecha, dóblelo por la mitad y póngalo encima de la mesa.                | 3                 |

|                                                                                                        |   |
|--------------------------------------------------------------------------------------------------------|---|
| Lea esta frase y haga lo que dice: "Cierre los ojos".                                                  | 1 |
| Escriba una frase (con sujeto y predicado).                                                            | 1 |
| Copia este dibujo<br> | 1 |

### ***III.2 ÍNDICE DE BARTHEL***

La escala de Barthel [43] mide la capacidad funcional de la persona para realizar de forma independiente las actividades básicas de la vida diaria (ABVD), como comer, trasladarse, asearse, ir al baño, caminar, vestirse y controlar esfínteres.

El profesional evalúa si el residente realiza cada actividad de forma independiente, con ayuda parcial o con dependencia total. Cada una de las actividades tiene una puntuación específica.

La puntuación sigue estos rangos (sobre 100 puntos):

- 0 a 20 puntos: Dependencia total.
- 21 a 35 puntos: Dependencia grave.
- 40 a 55 puntos: Dependencia moderada.
- 60 a 95 puntos: Dependencia leve.
- 100 puntos: Independencia total.

*Tabla 15 - Preguntas Índice de Barthel*

| <i>Actividad</i>          | <i>Opciones</i>                                                                                                               |
|---------------------------|-------------------------------------------------------------------------------------------------------------------------------|
| Comer                     | Independiente (10) / Necesita ayuda para cortar alimentos, extender mantequilla, etc. (5) / Dependiente (0)                   |
| Lavarse (ducha o baño)    | Independiente (5) / Dependiente (0)                                                                                           |
| Aseo personal             | Independiente (5) / Dependiente (0)                                                                                           |
| Vestirse                  | Independiente (10) / Necesita ayuda, pero realiza al menos la mitad de la tarea (5) / Dependiente (0)                         |
| Control vesical           | Continente (10) / Episodios ocasionales de incontinencia (5) / Incontinente (0)                                               |
| Uso del retrete           | Independiente (10) / Necesita algo de ayuda (5) / Dependiente (0)                                                             |
| Trasladarse (sillón-cama) | Independiente (15) / Mínima ayuda física o supervisión (10) / Gran ayuda, pero puede permanecer sentado (5) / Dependiente (0) |
| Deambulación              | Independiente (15) / Necesita ayuda o supervisión (10) / Independiente en silla de ruedas (5) / Dependiente (0)               |
| Subir y bajar escaleras   | Independiente (10) / Necesita ayuda o supervisión (5) / Dependiente (0)                                                       |
| Control intestinal        | Continente (10) / Episodios ocasionales de incontinencia (5) / Incontinente (0)                                               |

### ***III.3 ESCALA DE TINETTI***

La escala de Tinetti [44] mide la movilidad de la persona mayor mediante la observación directa del equilibrio (sentado y de pie) y las características de la marcha (longitud del paso, simetría, fluidez). Se utiliza para predecir el riesgo de caídas.

Para realizarlo, el profesional observa al residente realizar una serie de movimientos controlados, como levantarse de una silla, caminar unos metros o girar sobre sí mismo.

La puntuación sigue estos rangos (sobre 28 puntos, 16 de equilibrio y 12 de marcha):

- 0 a 18 puntos: Riesgo alto de caídas.
- 19 a 23 puntos: Riesgo moderado de caídas.
- 24 a 28 puntos: Riesgo bajo de caídas.

*Tabla 16 - Preguntas Escala de Tinetti para la evaluación del equilibrio*

| <i>Pregunta</i>                       | <i>Opciones</i>                                                                                                  |
|---------------------------------------|------------------------------------------------------------------------------------------------------------------|
| Equilibrio sentado                    | Se inclina o desliza en la silla (0) / Se mantiene seguro (1)                                                    |
| Intentos para levantarse              | Incapaz sin ayuda (0) / Capaz, pero necesita más de un intento (1) / Capaz de levantarse con un solo intento (2) |
| Levantarse                            | Imposible sin ayuda (0) / Capaz, pero usa los brazos para ayudarse (1) / Capaz sin usar los brazos (2)           |
| Equilibrio en bipedestación inmediata | Inestable (0) / Estable con apoyo (1) / Estable sin apoyo (2)                                                    |
| Equilibrio en bipedestación           | Inestable (0) / Estable con base amplia o apoyo (1) / Estable con base estrecha sin soporte (2)                  |

| <i>Pregunta</i>                               | <i>Opciones</i>                                                                                    |
|-----------------------------------------------|----------------------------------------------------------------------------------------------------|
| Empujar (test de empuje suave en el esternón) | Empieza a caerse (0) / Se tambalea, pero se mantiene (1) / Estable (2)                             |
| Ojos cerrados                                 | Inestable (0) / Estable (1)                                                                        |
| Giro de 360°                                  | Inestable o pasos discontinuos (0) / Estable o pasos continuos (1)                                 |
| Sentarse                                      | Inseguro o cae en la silla (0) / Usa brazos o movimiento brusco (1) / Seguro, movimiento suave (2) |

*Tabla 17 - Preguntas Escala de Tinetti para la evaluación de la marcha*

| <i>Pregunta</i>                            | <i>Opciones</i>                                                                      |
|--------------------------------------------|--------------------------------------------------------------------------------------|
| Inicio de la marcha                        | Vacila o necesita varios intentos (0) / Inicia la marcha sin vacilar (1)             |
| Longitud y altura del paso (pie derecho)   | No sobrepasa el pie izquierdo (0) / Sobrepasa el pie izquierdo (1)                   |
| Longitud y altura del paso (pie izquierdo) | No sobrepasa el pie derecho (0) / Sobrepasa el pie derecho (1)                       |
| Simetría del paso                          | Pasos asimétricos (0) / Pasos simétricos (1)                                         |
| Fluidez del paso                           | Paradas entre los pasos (0) / Pasos continuos (1)                                    |
| Trayectoria                                | Desviación marcada (0) / Desviación leve o usa ayudas (1) / Trayectoria recta (2)    |
| Postura al caminar                         | Talones separados o postura inestable (0) / Talones casi juntos, postura estable (1) |

### ***III.4 ESCALA DE NORTON***

La escala de Norton [45] mide el riesgo que tiene un paciente de presentar úlceras por presión (escaras). Evalúa cinco parámetros: estado físico general, estado mental, actividad, movilidad e incontinencia.

El profesional asigna una puntuación de 1 a 4 en cada dimensión según el estado del residente.

La puntuación sigue estos rangos (sobre 20 puntos):

- 5 a 11 puntos: Riesgo muy alto de padecer úlceras.
- 12 a 14 puntos: Riesgo evidente (Riesgo medio).
- Más de 14 puntos: Riesgo mínimo o sin riesgo.

*Tabla 18 - Preguntas escala de Norton*

| <i>Pregunta</i>  | <i>Opciones</i>                                                               |
|------------------|-------------------------------------------------------------------------------|
| Estado mental    | Orientado (4) / Algo confuso (3) / Respuesta inadecuada (2) / No responde (1) |
| Actividad        | Deambula (4) / Camina con ayuda (3) / Silla de ruedas (2) / Encamado (1)      |
| Movilidad actual | Completa (4) / Algo limitada (3) / Muy limitada (2) / Inmóvil (1)             |
| Incontinencia    | No (4) / A veces (3) / Orina (2) / Orina y heces (1)                          |
| Estado físico    | Bueno (4) / Insuficiente (3) / Malo (2) / Muy malo (1)                        |

### ***III.5 ESCALA DE DOWNTON***

La escala de Downton [46] es una herramienta de valoración diseñada para estimar el riesgo de caída en personas mayores institucionalizadas o con patología crónica. Se basa en la identificación de factores predisponentes que aumentan la probabilidad de sufrir caídas en el entorno residencial.

Evalúa cinco grandes bloques: antecedentes previos de caídas, medicación que pueda afectar al sistema nervioso central (sedantes, antihipertensivos, diuréticos, etc.), déficits sensoriales (alteraciones visuales o auditivas), estado mental (confusión o deterioro cognitivo) y tipo de deambulación (segura, con ayuda técnica o inestable).

Cada factor suma un punto cuando está presente, lo que permite una valoración rápida y objetiva del riesgo.

La puntuación sigue estos rangos:

- 0 a 2 puntos: Riesgo bajo de caída.
- 3 o más puntos: Riesgo alto de caída.

### ***III.6 TIMED UP & Go (TUG)***

La escala TUG [47] mide la agilidad y el equilibrio dinámico. Consiste en medir el tiempo o la dificultad de un residente para levantarse de una silla, caminar 3 metros y volver a sentarse. La puntuación (en tu sistema simplificado de 1 a 5) sigue estos rangos:

- 1 a 2: Riesgo normal de caídas (Movilidad aceptable).
- 3: Riesgo de caída (Marcha medianamente alterada).
- 4 a 5: Riesgo alto de caída (Marcha moderada o severamente alterada).

### ***III.7 ESCALA GDS (GLOBAL DETERIORATION SCALE)***

La escala GDS [48] mide el estadio de deterioro cognitivo en pacientes con enfermedades neurodegenerativas, dividiendo la evolución en siete etapas clínicas claras. La puntuación sigue estos rangos:

- GDS 1 a 3: Desde la normalidad hasta el deterioro cognitivo leve (olvidos benignos).
- GDS 4 a 5: Deterioro moderado (fases iniciales de la demencia; ayuda en tareas complejas).
- GDS 6 a 7: Deterioro grave a muy grave (pérdida de autonomía total y lenguaje).

*Tabla 19 - Descripción estados en la escala GDS*

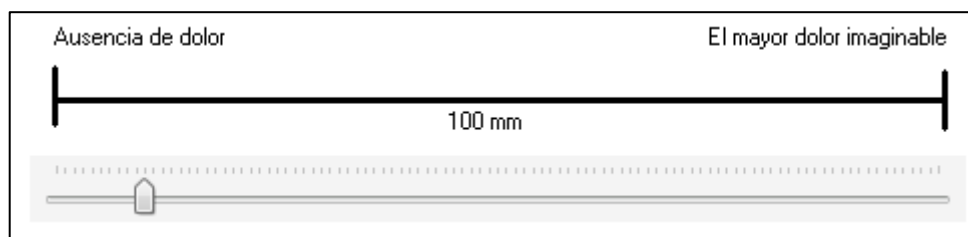
| <i>Descripción del estado cognitivo</i> | <i>Opciones</i>                                                                                         |
|-----------------------------------------|---------------------------------------------------------------------------------------------------------|
| Ausencia de deterioro cognitivo         | Sin problemas subjetivos / objetivos de memoria (1)                                                     |
| Deterioro cognitivo muy leve            | Quejas subjetivas de memoria (olvidar nombres o localización de objetos) (2)                            |
| Deterioro cognitivo leve                | Primeros déficits claros en memoria y concentración detectables en actividades sociales o laborales (3) |
| Deterioro cognitivo moderado            | Déficit en memoria reciente, dificultades para tareas complejas (dinero o viajar solo) (4)              |
| Deterioro cognitivo moderadamente grave | Necesita ayuda para algunas actividades diarias y presenta desorientación temporal o espacial (5)       |
| Deterioro cognitivo grave               | Pérdida significativa de memoria, cambios de personalidad y necesidad de ayuda en la vida diaria (6)    |
| Deterioro cognitivo muy grave           | Pérdida del lenguaje, dependencia total para las actividades básicas y alteraciones motoras (7)         |

### ***III.8 ESCALA EVA (ESCALA VISUAL ANALÓGICA)***

La escala EVA [49] mide la intensidad del dolor percibida por el residente. Es una herramienta subjetiva pero fundamental para correlacionar el malestar físico con la conducta.

La puntuación sigue estos rangos:

- 0 a 1: Ausencia o dolor muy leve.
- 2 a 6: Dolor moderado.
- 7 a 10: Dolor intenso a insoportable.



*Figura 13 - Escala EVA*

### ***III.9 ESCALA YESAVAGE (GDS-15)***

La escala Yesavage [50] mide la presencia de síntomas depresivos en la población mayor, centrándose más en el estado de ánimo y el abandono de actividades que en síntomas físicos.

Se trata de un cuestionario de 15 preguntas de respuesta dicotómica (sí/no).

La puntuación sigue estos rangos (versión corta de 15 ítems):

- 0 a 5 puntos: Normal (Sin síntomas depresivos).
- 6 a 9 puntos: Sugere de depresión (Depresión leve).
- 10 o más puntos: Indicativo de depresión (Depresión establecida).

*Tabla 20 - Preguntas escala Yesavage*

| <i>Pregunta</i>                                                           | <i>Opciones (Puntuación)</i> |
|---------------------------------------------------------------------------|------------------------------|
| ¿Está básicamente satisfecho con su vida?                                 | Sí (1) / No (0)              |
| ¿Ha renunciado a muchas de sus actividades e intereses?                   | Sí (1) / No (0)              |
| ¿Siente que su vida está vacía?                                           | Sí (1) / No (0)              |
| ¿Se encuentra a menudo aburrido?                                          | Sí (1) / No (0)              |
| ¿Tiene a menudo buenos ánimos?                                            | Sí (1) / No (0)              |
| ¿Tiene miedo de que algo malo le esté pasando?                            | Sí (1) / No (0)              |
| ¿Se siente feliz muchas veces?                                            | Sí (1) / No (0)              |
| ¿Se siente a menudo abandonado/a?                                         | Sí (1) / No (0)              |
| ¿Prefiere quedarse en casa en lugar de salir y hacer cosas nuevas?        | Sí (1) / No (0)              |
| ¿Encuentra que tiene más problemas de memoria que la mayoría de la gente? | Sí (1) / No (0)              |
| ¿Piensa que es maravilloso vivir?                                         | Sí (1) / No (0)              |
| ¿Le resulta difícil iniciar nuevos proyectos?                             | Sí (1) / No (0)              |
| ¿Se siente lleno/a de energía?                                            | Sí (1) / No (0)              |
| ¿Siente que su situación es desesperada?                                  | Sí (1) / No (0)              |
| ¿Cree que mucha gente está mejor que usted?                               | Sí (1) / No (0)              |

### III.10 INVENTARIO NEUROPSIQUIÁTRICO (NPI)

La escala NPI [51] mide la frecuencia y la gravedad de alteraciones conductuales como agitación, agresión, apatía, desinhibición o trastornos del sueño. La puntuación puede ser:

- Puntuaciones bajas: Ausencia de trastornos de conducta significativos.
- Puntuaciones altas: Presencia de síntomas neuropsiquiátricos graves que requieren intervención clínica urgente.

Tabla 21 - Síntomas evaluados en NPI

| Síntoma evaluado                        | Opciones                          |
|-----------------------------------------|-----------------------------------|
| Delirios                                | Frecuencia (0–4) × Gravedad (1–3) |
| Alucinaciones                           | Frecuencia (0–4) × Gravedad (1–3) |
| Agitación o agresividad                 | Frecuencia (0–4) × Gravedad (1–3) |
| Depresión o disforia                    | Frecuencia (0–4) × Gravedad (1–3) |
| Ansiedad                                | Frecuencia (0–4) × Gravedad (1–3) |
| Euforia o júbilo                        | Frecuencia (0–4) × Gravedad (1–3) |
| Apatía o indiferencia                   | Frecuencia (0–4) × Gravedad (1–3) |
| Desinhibición                           | Frecuencia (0–4) × Gravedad (1–3) |
| Irritabilidad o inestabilidad emocional | Frecuencia (0–4) × Gravedad (1–3) |
| Conducta motora aberrante               | Frecuencia (0–4) × Gravedad (1–3) |
| Alteraciones del sueño                  | Frecuencia (0–4) × Gravedad (1–3) |
| Alteraciones del apetito o alimentación | Frecuencia (0–4) × Gravedad (1–3) |

### ***III.11 ESCALA MINI NUTRITIONAL ASSESSMENT (MNA)***

El Mini Nutritional Assessment (MNA) [52] es una herramienta validada para la detección precoz del riesgo de desnutrición en personas mayores. Se utiliza ampliamente en residencias y entornos sociosanitarios debido a su sensibilidad para identificar tanto desnutrición establecida como riesgo nutricional.

La versión completa de la escala consta de 18 ítems, agrupados en cuatro áreas principales de evaluación: evaluación antropométrica (peso, talla, índice de masa y otras medidas corporales), evaluación global (estilo de vida, movilidad, presencia de enfermedades y medicación), evaluación dietética (ingesta alimentaria, número de comidas diarias y consumo de determinados grupos de alimentos) y autoevaluación del estado de salud y del estado nutricional.

Además de la versión completa, existe una versión abreviada de cribado inicial denominada MNA-SF (Mini Nutritional Assessment – Short Form) compuesta por 6 preguntas, que permite realizar una detección rápida del riesgo nutricional. En caso de que el resultado indique posible riesgo de malnutrición, se recomienda completar la evaluación utilizando la versión completa de la escala.

La puntuación máxima en la versión completa es de 30 puntos. Los rangos de interpretación son:

- 24 a 30 puntos: Estado nutricional normal.
- 17 a 23,5 puntos: Riesgo de desnutrición.
- Menos de 17 puntos: Desnutrición establecida.

*Tabla 22 - Preguntas escala de Mini Nutritional Assessment I*

| <i>Item</i> | <i>Pregunta</i>                                                                                                                                       | <i>Opciones (Puntuación)</i>                                                                            |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| A           | ¿Ha perdido el apetito o ha comido menos por falta de apetito, problemas digestivos o dificultades de masticación o deglución en los últimos 3 meses? | Ha comido mucho menos (0) / Ha comido menos (1) / Ha comido igual (2)                                   |
| B           | Pérdida reciente de peso (últimos 3 meses)                                                                                                            | >3 kg (0) / No lo sabe (1) / Entre 1 y 3 kg (2) / No ha habido pérdida de peso (3)                      |
| C           | Movilidad                                                                                                                                             | De la cama al sillón (0) / Autonomía en el interior (1) / Sale del domicilio (2)                        |
| D           | ¿Ha tenido una enfermedad aguda o situación de estrés psicológico en los últimos 3 meses?                                                             | Sí (0) / No (2)                                                                                         |
| E           | Problemas neuropsicológicos                                                                                                                           | Demencia o depresión grave (0) / Demencia leve (1) / Sin problemas psicológicos (2)                     |
| F           | Índice de masa corporal (IMC)                                                                                                                         | IMC < 19 (0) / $19 \leq \text{IMC} < 21$ (1) / $21 \leq \text{IMC} < 23$ (2) / $\text{IMC} \geq 23$ (3) |

Tabla 23 - Preguntas escala de Mini Nutritional Assessment II

| Item | Pregunta                                                            | Opciones (Puntuación)                                                                            |
|------|---------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| G    | ¿Vive independiente en su domicilio?                                | Sí (0) / No (1)                                                                                  |
| H    | ¿Más de 3 medicamentos/día?                                         | Sí (0) / No (1)                                                                                  |
| I    | ¿Presenta úlceras o lesiones cutáneas?                              | Sí (0) / No (1)                                                                                  |
| J    | ¿Cuántas comidas completas/día?                                     | 1 comida (0) / 2 comidas (1) / 3 comidas (2)                                                     |
| K    | Consumo de proteínas (lácteos, huevos o legumbres, carne o pescado) | 0 o 1 respuestas afirmativas (0) / 2 respuestas afirmativas (0,5) / 3 respuestas afirmativas (1) |
| L    | ¿Consume frutas o verduras al menos dos veces al día?               | No (0) / Sí (1)                                                                                  |
| M    | ¿Consume frutas o verduras al menos dos veces al día?               | Menos de 3 vasos (0) / De 3 a 5 vasos (0,5) / Más de 5 vasos (1)                                 |
| N    | Forma de alimentarse                                                | Necesita ayuda (0) / Solo con dificultad (1) / Solo sin dificultad (2)                           |
| O    | ¿Se considera el paciente bien nutrido?                             | Malnutrición grave (0) / No lo sabe o malnutrición moderada (1) / Sin problemas (2)              |
| P    | ¿Se considera bien nutrido?                                         | Peor (0) / No sabe (0,5) / Igual (1) / Mejor (2)                                                 |
| Q    | Circunferencia braquial (CB)                                        | CB < 21 cm (0) / $21 \leq CB \leq 22$ cm (0,5) / CB > 22 cm (1)                                  |
| R    | Circunferencia de la pantorrilla (CP)                               | CP < 31 cm (0) / CP $\geq 31$ cm (1)                                                             |

### ***III.12 ESCALA PAINAD (PAIN ASSESSMENT IN ADVANCED DEMENTIA)***

La escala PAINAD (Pain Assessment in Advanced Dementia) [53] es una herramienta diseñada para valorar el dolor en personas con deterioro cognitivo moderado o grave que presentan dificultades de comunicación verbal. En estos casos, la escala EVA puede no ser fiable, por lo que se recurre a métodos observacionales.

La escala evalúa cinco parámetros conductuales:

- Respiración (normal o alterada).
- Vocalización negativa (quejidos, gemidos).
- Expresión facial (gestos de dolor).
- Lenguaje corporal (rigidez, agitación, defensa).
- Capacidad de alivio o consolabilidad.

Cada uno de estos ítems se puntúa de 0 a 2, obteniéndose una puntuación total entre 0 y 10 puntos. La interpretación es equivalente a la escala EVA:

- 0 puntos: Ausencia de dolor.
- 1 a 3 puntos: Dolor leve.
- 4 a 6 puntos: Dolor moderado.
- 7 a 10 puntos: Dolor intenso.