



Nº Expediente: 00001-00113663

Solicitante: [REDACTED]

NIF: [REDACTED]

Asunto: Resolución en materia de acceso a la información presentada al amparo del artículo 12 de la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno (en adelante también denominada "LTAIBG").

En respuesta a la solicitud presentada por D. [REDACTED], recibida ante la Unidad de Información de Transparencia de la Seguridad Social el 27 de enero de 2026, y en virtud de la cual solicita información relativa a:

Que, en ejercicio del derecho reconocido en el artículo 12 de la Ley 19-2013, y tras analizar la documentación técnica facilitada por esa Administración en los expedientes 00001-00105625 y 00001-00107257, relativos al modelo predictivo de Gradient Boosting empleado para la priorización de "Primeras Citas" en procesos de Incapacidad Temporal, se han identificado contradicciones técnicas y vacíos de información que hacen necesario solicitar aclaraciones adicionales, en atención al evidente interés público del sistema descrito.

En particular, se observa una divergencia significativa entre la descripción funcional del modelo y los resultados reflejados en los documentos de Feature Importance, así como el uso de variables sensibles (sexo y localización geográfica) sin que conste evaluación pública suficiente de su impacto.

POR TODO ELLO, SOLICITA

Que se facilite el acceso a la siguiente información pública:

1. Desglose completo del "Target Encoding" aplicado a la variable provincial

Dado que la variable DS_PROVINCIA_CORTA figura como la de mayor relevancia del modelo (peso 1.0) y que se le aplica una transformación de tipo target encoding, se solicita el listado íntegro de los valores numéricos resultantes de dicha codificación para cada provincia, indicando la probabilidad o puntuación asociada a la finalización del proceso de Incapacidad Temporal que el sistema asigna a cada territorio.

2. Justificación técnica de la jerarquía real de variables del modelo

Se solicita explicación técnica detallada que aclare por qué la variable de localización geográfica presenta una relevancia aproximadamente 30 veces superior a la variable SPS_MEDIAN (duración mediana de la baja, peso 0,034), cuando en la resolución administrativa correspondiente se identifica la duración estimada como el principal criterio de priorización.

3. Tabla de valores de los indicadores de agravante por diagnóstico

Se solicita la tabla o matrices de cálculo utilizadas para los indicadores IND_AGRAVANTE_SEXO e IND_AGRAVANTE_RECAIDA, especificando los valores aplicados por capítulos diagnósticos (CIE-10), a fin de conocer qué patologías y perfiles son considerados de mayor o menor probabilidad de alta por el sistema.

4. Informes de seguimiento y evaluación del modelo en la plataforma Microstrategy

Conforme a la documentación funcional del sistema, se solicita copia de los informes mensuales de evaluación del rendimiento del modelo elaborados en la plataforma Microstrategy durante los últimos 12 meses disponibles, incluyendo métricas de acierto, discrepancias y resultados reales frente a predicciones.

5. Evaluaciones internas sobre impacto y sesgo algorítmico

Dado el uso explícito de la variable sexo y de indicadores derivados de esta, y ante la inexistencia de auditorías externas reconocidas, se solicita cualquier informe interno, nota técnica o evaluación de impacto destinada a verificar la ausencia de discriminación por razón de sexo en la priorización de citas médicas.

6. Justificación técnica de la vigencia de los datos de entrenamiento (noviembre de 2020)

Se solicita el informe técnico o criterio documentado que sustenta la decisión de mantener en producción un modelo entrenado con datos cuya última actualización data de noviembre de 2020, especialmente considerando los cambios sanitarios, laborales y normativos posteriores.

La Gerencia de Informática de la Seguridad Social (GISS), como órgano competente en su ámbito, adopta la siguiente

RESOLUCIÓN:

Una vez analizada la solicitud y en referencia a las cuestiones que son competencia de esta Gerencia de Informática de la Seguridad Social se resuelve admitir la solicitud reconociendo el derecho **al acceso parcial** de la información a que se refiere la solicitud, en los términos establecidos en el artículo 105, letra b) de la Constitución, y en los artículos 12 y siguientes de la Ley 19/2013, de 9 de diciembre. A tal efecto, y en las cuestiones que son del ámbito de esta Gerencia de Informática de la Seguridad Social se informa de lo siguiente:

RESPUESTA:

1. Desglose completo del “Target Encoding” aplicado a la variable provincial

Dado que la variable DS_PROVINCIA_CORTA figura como la de mayor relevancia del modelo (peso 1.0) y que se le aplica una transformación de tipo target encoding, se solicita el listado íntegro de los valores numéricos resultantes de dicha codificación para cada provincia, indicando la probabilidad o puntuación asociada a la finalización del proceso de Incapacidad Temporal que el sistema asigna a cada territorio.

El modelo **no asigna a ningún territorio una probabilidad específica o puntuación asociada a la finalización del proceso de incapacidad temporal**, dado que la variable territorial forma parte de un enfoque **multivariante** en el que las combinaciones de factores aportan información conjunta, sin generar indicadores individualizados por provincia.

2. Justificación técnica de la jerarquía real de variables del modelo

Se solicita explicación técnica detallada que aclare por qué la variable de localización geográfica presenta una relevancia aproximadamente 30 veces superior a la variable SPS_MEDIAN (duración mediana de la baja, peso 0,034), cuando en la resolución administrativa correspondiente se identifica la duración estimada como el principal criterio de priorización.

Es destacable que el documento Feature Importance compartido en la resolución anterior refleja la importancia relativa de las variables. Por tanto, el valor 1,0 asignado a DS_PROVINCIA_CORTA no indica que esta variable sea treinta veces más determinante en términos sustantivos que SPS_MEDIAN, sino que es la variable con mayor peso relativo dentro del conjunto.

El modelo utiliza técnicas multivariantes en las que las variables interaccionan entre sí. En este contexto, la variable geográfica puede reflejar **diferencias organizativas y operativas** entre direcciones provinciales, lo cual explica que aporte información útil en el modelo.

Sin embargo, esta contribución **no debe interpretarse como un criterio de priorización individual**, ni como una ponderación aislada, ya que el modelo evalúa conjuntos de variables en interacción y la duración estimada de la IT.

3. Tabla de valores de los indicadores de agravante por diagnóstico

Se solicita la tabla o matrices de cálculo utilizadas para los indicadores IND_AGRAVANTE_SEXO e IND_AGRAVANTE_RECAIDA, especificando los valores aplicados por capítulos diagnósticos (CIE-10), a fin de conocer qué patologías y perfiles son considerados de mayor o menor probabilidad de alta por el sistema.

Estos indicadores se construyen a partir de **datos agregados y anonimizados**, y **no asignan probabilidades de alta médica** a diagnósticos individuales. Su función es únicamente contribuir al análisis global, sin generar etiquetas individualizadas ni deterministas.

Conforme a los principios de protección del modelo y seguridad operacional establecidos en las directrices de gobernanza interna, no se facilitan valores específicos pues ello permitiría inferir la lógica interna de ponderación del modelo, facilitando su ingeniería inversa y su posible manipulación deliberada.

Los valores desagregados por diagnóstico podrían combinarse con otra información disponible para extraer conclusiones sobre el comportamiento del modelo ante perfiles clínicos específicos, comprometiendo la integridad operativa del sistema.

La información relativa a patrones de decisión asociados a categorías diagnósticas concretas tiene la consideración de información técnica sensible por tanto no se facilitan, en aplicación del artículo 14-1k de la Ley 19/2013 que establece límites en el acceso a la información por la garantía de la confidencialidad o el secreto requerido en procesos de toma de decisión.

4. Informes de seguimiento y evaluación del modelo en la plataforma Microstrategy

Conforme a la documentación funcional del sistema, se solicita copia de los informes mensuales de evaluación del rendimiento del modelo elaborados en la plataforma Microstrategy durante los últimos 12 meses disponibles, incluyendo métricas de acierto, discrepancias y resultados reales frente a predicciones.

En la resolución del expediente 00001-00105625 se facilitaron datos estadísticos agregados sobre el funcionamiento del modelo durante el primer semestre de 2025.

Cabe añadir que se realiza un seguimiento periódico del rendimiento del modelo. Este seguimiento tiene como finalidad detectar posibles degradaciones del modelo, verificar la coherencia de las predicciones con los resultados reales y fundamentar las decisiones sobre mantenimiento o actualización del sistema.

Los informes internos mensuales de rendimiento incluyen métricas y datos con **información restringida**, debido a:

- presencia de métricas detalladas,
- posibles patrones explotables,
- información operativa interna.

No se facilitan informes técnicos internos completos que documentan este seguimiento, dado que contienen información operativa sensible cuya divulgación podría facilitar la manipulación del sistema. Por otro lado, se recuerda que la Ley 19/2013 en el artículo 14-1g establece límites al derecho de acceso a la información pública en lo referente a funciones de vigilancia, inspección y control.

5. Evaluaciones internas sobre impacto y sesgo algorítmico

Dado el uso explícito de la variable sexo y de indicadores derivados de esta, y ante la inexistencia de auditorías externas reconocidas, se solicita cualquier informe interno, nota técnica o evaluación de impacto destinada a verificar la ausencia de discriminación por razón de sexo en la priorización de citas médicas.

En la resolución del expediente 00001-00103694 se indicó que se realizó una evaluación de impacto conforme al Reglamento (UE) 2024/1689 y que el sistema no incurre en ninguno de los usos prohibidos contemplados en dicha norma. En respuesta a la solicitud de mayor detalle sobre el análisis de sesgo por razón de sexo, se informa que el sistema incorpora mecanismos de supervisión orientados a garantizar el cumplimiento de los principios de **no discriminación**, uso responsable de datos y **equidad algorítmica**, conforme al marco normativo aplicable.

6. Justificación técnica de la vigencia de los datos de entrenamiento (noviembre de 2020)

Se solicita el informe técnico o criterio documentado que sustenta la decisión de mantener en producción un modelo entrenado con datos cuya última actualización data de noviembre de 2020, especialmente considerando los cambios sanitarios, laborales y normativos posteriores.

En las resoluciones anteriores requeridas por el solicitante se ha informado que el modelo en producción fue actualizado por última vez en noviembre de 2020.

En la actualidad se mantiene en producción tras un proceso continuo de **monitorización periódica del rendimiento** por parte del órgano competente. Según la información operativa disponible, se revisan de forma periódica los **resultados** obtenidos, comparando:

- expedientes priorizados mediante el modelo,
- expedientes gestionados sin modelo,
- otros criterios complementarios.

Los análisis realizados muestran que el uso del modelo **aporta mejoras operativas significativas**, y no se han identificado degradaciones que justifiquen un reentrenamiento inmediato.

Contra la presente resolución que pone fin a la vía administrativa, y sin perjuicio de otras acciones, de conformidad con lo dispuesto por el artículo 88.3 de la Ley 39/2015, y los artículos 6 y siguientes de la Ley 29/1998, de 13 de julio, reguladora de la Jurisdicción Contencioso-administrativa, podrá interponer recurso ante el Juzgado Central de lo Contencioso-administrativo que por turno corresponda, en el plazo de dos



MINISTERIO
DE INCLUSIÓN,
SEGURIDAD SOCIAL
Y MIGRACIONES

meses, o, previa y potestativamente, reclamación ante el Consejo de Transparencia y Buen Gobierno en el plazo de un mes; en ambos casos el plazo se contará desde el día siguiente a aquel durante el cual se practique la notificación de la presente resolución.

Madrid, a la fecha de la firma electrónica
EL GERENTE DE INFORMÁTICA

Firmado electronicamente por: ESCUDERO RIVAS
CARLOS
26.03.2026 14:12:57 CET