

PROVINCIA DE ENTRE RÍOS

CONSEJO FEDERAL DE INVERSIONES

PROYECTO "ANÁLISIS INTEGRADO DE DATA Y COMUNICACIÓN DIGITAL AVANZADA  
ETAPA 2"

EX-2025-00023271- -CFI-GES#DC

Informe Final

DICIEMBRE 2025.

CONSULTORA: Digital 360 S.R.L.

## Indice

i.	Resumen de proyecto.....	3
ii.	Principales logros.....	4
<b>iii.</b>	<b>Informe tecnico de tareas:.....</b>	<b>5</b>
<b>iv.</b>	<b>Tarea 1: Seguimiento continuo de métricas y asesoramiento estratégico.....</b>	<b>6</b>
v.	1.1. Monitoreo y análisis de métricas clave:.....	7
vi.	1.2 Desarrollo y optimización de reportes automatizados.....	8
vii.	1.3. Asesoramiento estratégico en la toma de decisiones:.....	9
<b>viii.</b>	<b>Tarea 2: Integración de nuevas fuentes de datos en el sistema de monitoreo..</b>	<b>10</b>
ix.	2.1 Incorporación de otros proyectos CFI y otras áreas de gobierno.....	11
x.	2.2. Unificación y centralización de la información en un único tablero de control:...	12
xi.	2.3. Optimización del procesamiento y almacenamiento de datos:.....	13
<b>xii.</b>	<b>Tarea 3: Optimización de la plataforma de escucha activa y tendencias.....</b>	<b>14</b>
xiii.	3.1 Nuevas capacidades de Machine Learning.....	14
xiv.	3.2. Ampliación de la cobertura de escucha activa:.....	15
xv.	3.3. Mejora de la visualización de tendencias y segmentación de datos:.....	15
<b>xvi.</b>	<b>Tarea 4: Capacitación y transferencia de conocimiento.....</b>	<b>16</b>
xvii.	4.1 Expansión del programa de formación a nuevos equipos.....	16
xviii.	4.2 Desarrollo de módulos avanzados de capacitación.....	17
<b>xix.</b>	<b>Conclusiones.....</b>	<b>20</b>
<b>xx.</b>	<b>Anexo I - Descripción de arquitectura técnica.....</b>	<b>21</b>
<b>xxi.</b>	<b>Anexo II - Descripción de fuentes de datos.....</b>	<b>22</b>
<b>xxii.</b>	<b>Anexo III - Conceptos fundamentales de los procesos de transferencia.....</b>	<b>26</b>

## Resumen proyecto

El presente informe final consolida el trabajo realizado en el marco del proyecto “Análisis Integrado de Data y Comunicación Digital Avanzada – Etapa 2”, orientado a fortalecer el monitoreo y análisis de la estrategia de comunicación digital del Gobierno de Entre Ríos mediante la integración de fuentes de datos, la estandarización de procesos de actualización y la disponibilidad de tableros de control para seguimiento operativo.

Durante la ejecución se avanzó sobre un esquema de centralización en una única plataforma web, incorporando un backend relacional como fuente de verdad para los indicadores y desarrollando pipelines de extracción y normalización en Python para unificar estructuras heterogéneas provenientes de distintas herramientas.

Sobre esa base, se implementaron módulos de visualización y tableros orientados a lectura temporal, segmentación y trazabilidad, con foco en habilitar consultas consistentes y comparables.

### **Las tareas ejecutadas se organizaron en las tres líneas principales:**

1. **Seguimiento continuo de métricas y asesoramiento estratégico:** incorporación de variables y métricas de trazabilidad temporal, y construcción de vistas de monitoreo para redes, web, CRM y otros módulos relevantes.
2. **Integración de nuevas fuentes y centralización:** desarrollo de conectores y procesos de actualización (automáticos y on demand), normalización y persistencia de datos en un repositorio relacional único para alimentar todos los tableros.
3. **Optimización de escucha activa y tendencias:** ampliación de cobertura de fuentes, y mejoras en visualizaciones interactivas y segmentación para facilitar la exploración de indicadores y patrones.

El resultado es una base funcional para el monitoreo integrado de la comunicación, con componentes reutilizables de ingesta, almacenamiento y presentación. La plataforma consolida indicadores de distintos dominios en un repositorio común y expone tableros que permiten seguimiento por periodo, cortes básicos de segmentación y trazabilidad de actualizaciones. Sobre esta base, las próximas iteraciones pueden incorporar nuevas fuentes, ampliar el set de métricas y profundizar análisis sin modificar la lógica central de integración y consulta.

## Principales logros

### **1. Consolidación de una plataforma única de consulta**

Se consolidó una plataforma web única para consulta de indicadores, con módulos diferenciados y una base común de datos. Esto reduce la dependencia de consultas manuales y facilita el acceso consistente a métricas por parte de los equipos de trabajo.

### **2. Integración y automatización de fuentes de datos**

Se avanzó en la integración de fuentes existentes y nuevas mediante pipelines de extracción y actualización. Esto incluye fuentes que ya contaban con actualización automática y fuentes que se consultan manualmente y fueron incorporadas a procesos reproducibles, con posibilidad de ejecución programada u on demand.

### **3. Centralización y normalización en backend relacional**

Se centralizó la información en un repositorio relacional único y se normalizaron estructuras para asegurar consistencia entre módulos. La manipulación y consulta se realiza a través de modelos y ORM, lo que ordena el acceso a datos y facilita la construcción de vistas agregadas sin depender de SQL artesanal.

### **4. Métricas con trazabilidad temporal y capacidad de seguimiento**

Se incorporaron métricas y recortes temporales que permiten seguimiento por periodo y comparación de evolución, mejorando la trazabilidad de acciones de comunicación y el soporte a análisis operativo.

### **5. Visualizaciones y segmentación para exploración de indicadores**

Se incorporaron tableros y visualizaciones interactivas orientadas a lectura temporal y segmentación, incluyendo filtros y componentes reutilizables que facilitan la exploración de la información por diferentes dimensiones (módulo, periodo y categorías principales).

## Links

Acceso a servidor Dev	User	Pass
<a href="#">Link</a>	admin	admin

# Informe técnico final

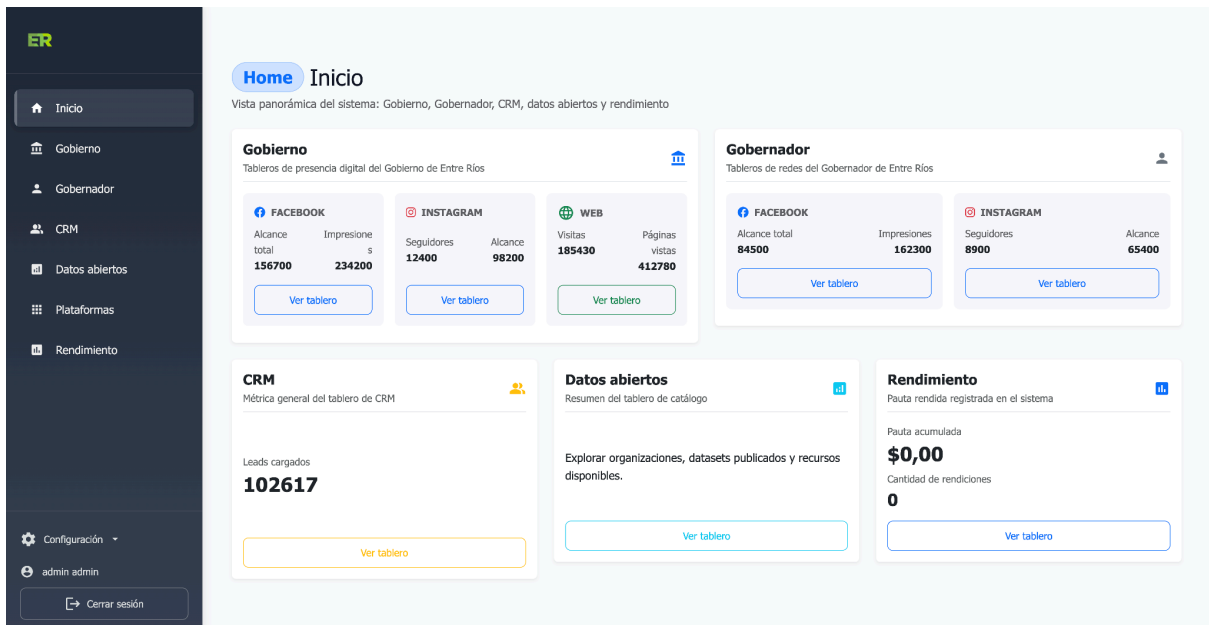
## Tarea 1: Seguimiento continuo de métricas y asesoramiento estratégico.

### 1.1. Monitoreo y análisis de métricas clave:

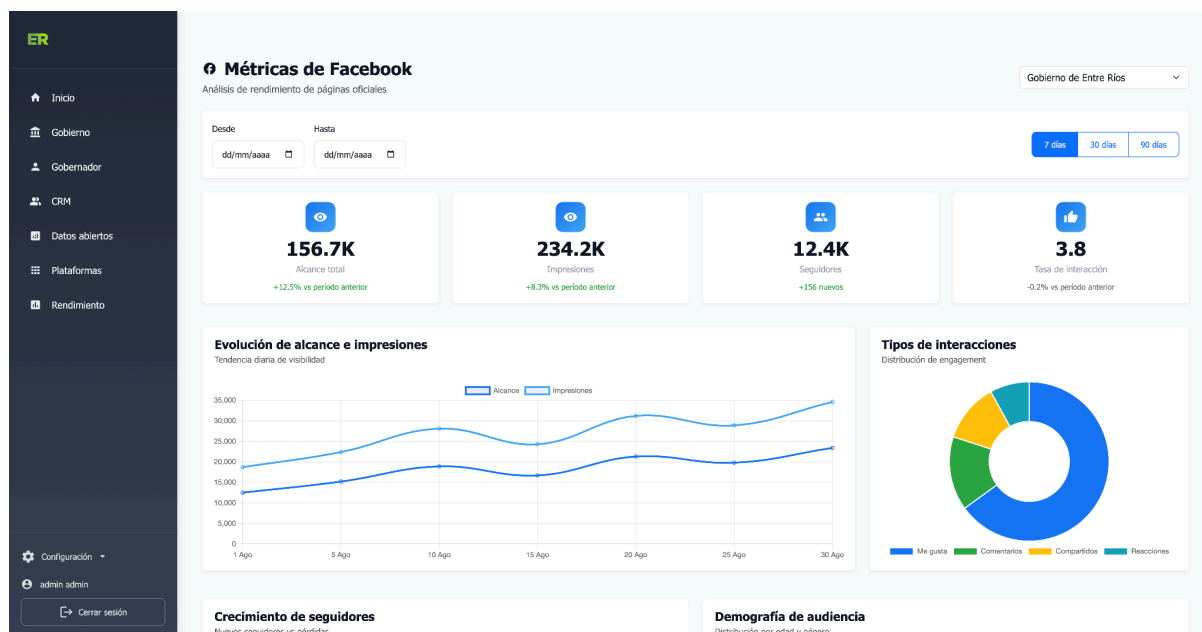
Al inicio del proyecto, el seguimiento de métricas se encontraba resuelto principalmente a través de tableros en Looker Studio y consultas complementarias, con un alcance acotado y sin una integración sistemática entre fuentes. En ese contexto, una parte de los indicadores se revisaba con recortes parciales y dependía de procesos manuales para consolidar información y generar comparaciones.

El trabajo realizado en este eje se orientó a ampliar el set de métricas disponibles y a estandarizar su lectura dentro de una plataforma única, incorporando trazabilidad temporal y criterios consistentes de consulta. Esto incluyó la incorporación de variables que no estaban representadas en los tableros previos y la organización de recortes por periodo para facilitar comparaciones y seguimiento operativo.

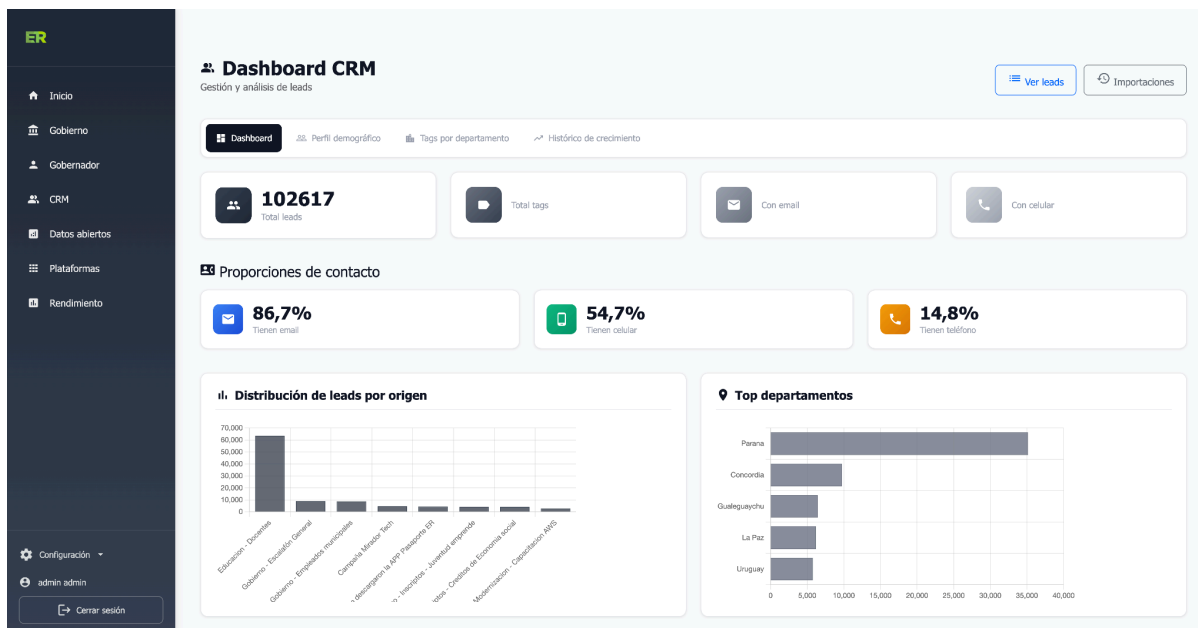
Estos avances se evidencian en el tablero de inicio con el resumen general y en los módulos específicos de métricas por canal, donde se incorporaron indicadores y recortes temporales para lectura comparativa.



En términos de integración, se fortaleció la conexión con el CRM y la extracción de información asociada a campañas y mailing, de modo de disponibilizar métricas de volumen y evolución junto con dimensiones de segmentación relevantes para el análisis (por ejemplo, tags y atributos de campaña). En paralelo, se incorporaron indicadores vinculados a recursos de datos abiertos y su estado de actualización, ampliando el monitoreo hacia variables de cumplimiento y calidad de publicación que no se encontraban contempladas en el esquema original.



En la plataforma, esto se refleja en el módulo de CRM, donde se visualizan métricas generales, segmentaciones y vistas históricas, y en el módulo de datos abiertos, donde se expone el estado de publicación, metadatos y elementos de seguimiento asociados a datasets y recursos.



Como criterio general, se priorizó que las métricas queden definidas de forma consistente entre módulos, evitando diferencias de cálculo según el origen o el tablero donde se consulten. Para esto, se organizaron procesos de extracción y normalización que sostienen un esquema de actualización repetible, reducen la intervención manual en la preparación de datos y permiten que los indicadores se consulten con el mismo criterio en diferentes vistas.

Este conjunto de cambios transforma un esquema centrado en visualizaciones aisladas hacia una plataforma integral de gestión y centralización de fuentes, métricas y tableros, con un monitoreo más integrado, métricas comparables y trazables en el tiempo. A su vez, deja una base estable para sumar progresivamente nuevos indicadores y vistas de análisis sin modificar la lógica central de consulta.

## 1.2. Desarrollo y automatización de reportes avanzados:

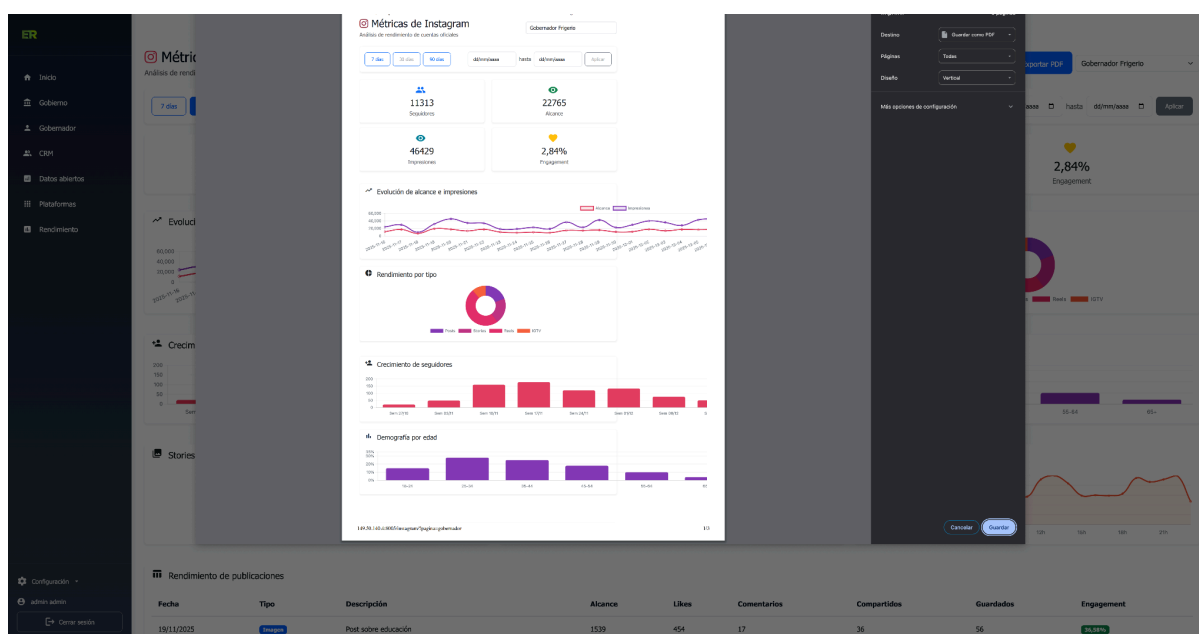
En paralelo al monitoreo en pantalla, se trabajó en habilitar mecanismos de salida que permitan convertir recortes de consulta en reportes compartibles para uso operativo. El punto de partida era un esquema donde los tableros servían como consulta, pero la documentación de hallazgos y la circulación interna dependían de capturas puntuales o de reconstrucciones manuales del estado de cada pantalla.

Para resolver esto, se incorporó una modalidad de exportación manual a PDF basada en la impresión de pantalla, priorizando que el reporte respete exactamente el recorte aplicado (filtros, periodo y selección de cuenta) y que sea reproducible bajo un criterio homogéneo de presentación. Este enfoque permite generar salidas rápidas para

adjuntar en instancias de seguimiento, sin introducir nuevas dependencias de procesamiento ni pipelines adicionales.



En la plataforma, la funcionalidad se evidencia mediante un botón de “Exportar PDF” en pantallas de consulta representativas. En el módulo de CRM, se habilita en el listado de leads y en el historial de importaciones, de modo de documentar tanto recortes de segmentación como el registro de ingestas y su trazabilidad. En datos abiertos, se incorpora tanto en la pantalla de exploración (listado de datasets con filtros) como en el detalle de un dataset, donde el PDF funciona como ficha de consulta con metadatos y recursos. En redes sociales, la exportación se encuentra disponible en los tableros de métricas de Facebook e Instagram, permitiendo documentar lecturas por cuenta y ventana temporal, con KPIs, evolución y desagregaciones visibles.



De este modo, el sistema incorpora una capa de reporte avanzada orientada a operación: convertir una consulta en un documento exportable sin perder contexto, manteniendo consistencia entre módulos y reduciendo tiempos de armado manual de reportes.

### 1.3. Asesoramiento estratégico en la toma de decisiones:

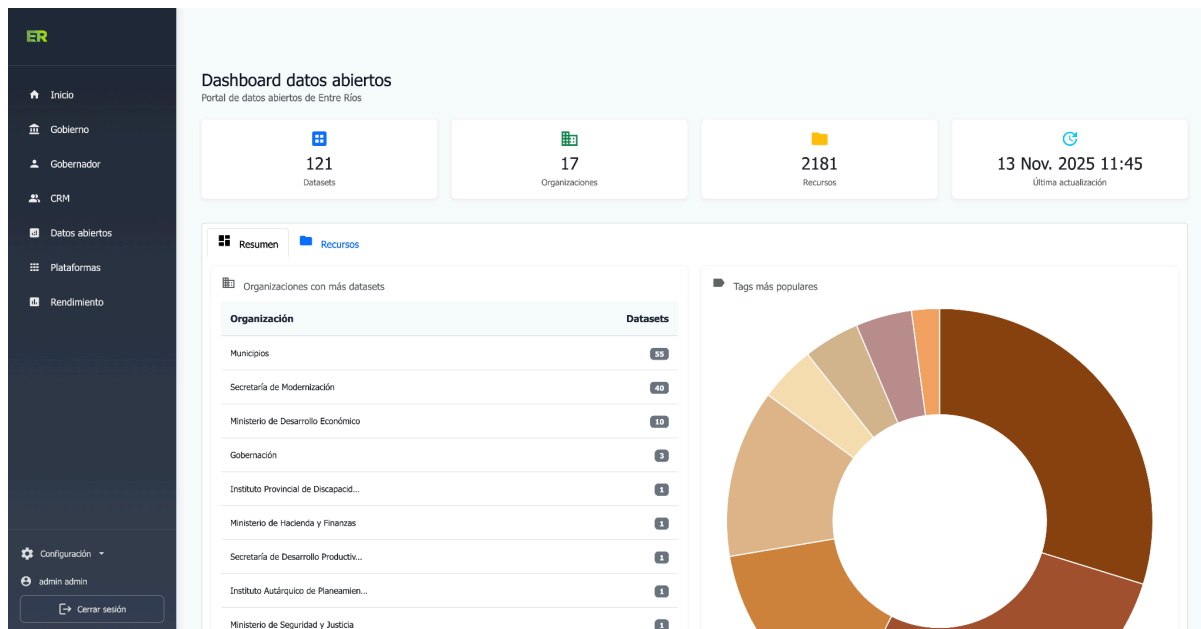
Sobre la base del monitoreo ampliado y la integración de fuentes, el proyecto incorporó una capa de análisis orientada a apoyar decisiones de comunicación con información consolidada. En particular, se construyeron métricas analíticas compuestas a partir de métricas primarias disponibles en los distintos canales (por ejemplo, alcance e

interacción en redes, volumen y segmentación de leads en CRM, rendimiento de campañas y variables de consumo de información), con el objetivo de permitir lecturas comparables entre periodos y entre módulos.

En esta línea, se incorporó trazabilidad temporal para que las decisiones no se basen únicamente en valores puntuales, sino en la evolución de los indicadores. Esto habilita evaluar cambios asociados a acciones concretas (ajustes de pauta, cambios de contenido, campañas específicas), y diferenciar variaciones estructurales de oscilaciones de corto plazo.

Un componente relevante del asesoramiento fue la incorporación de indicadores vinculados a datos abiertos, que permiten monitorear el estado de actualización y robustez de los recursos publicados. A partir de estas métricas, se fortaleció la recomendación sobre política de actualización y calidad de publicación (frecuencia, completitud y formatos), integrando este eje dentro del mismo marco de monitoreo de comunicación.

Estos elementos se evidencian en el tablero de inicio (lectura integrada de indicadores), en los módulos de métricas por canal (lectura comparativa y evolución), en el módulo de CRM (segmentación, histórico y lectura por tags) y en el módulo de datos abiertos (estado de publicación, metadatos y seguimiento de recursos). En conjunto, la plataforma permite sostener un circuito de revisión y ajuste con información trazable, reduciendo la dependencia de interpretaciones aisladas por fuente.



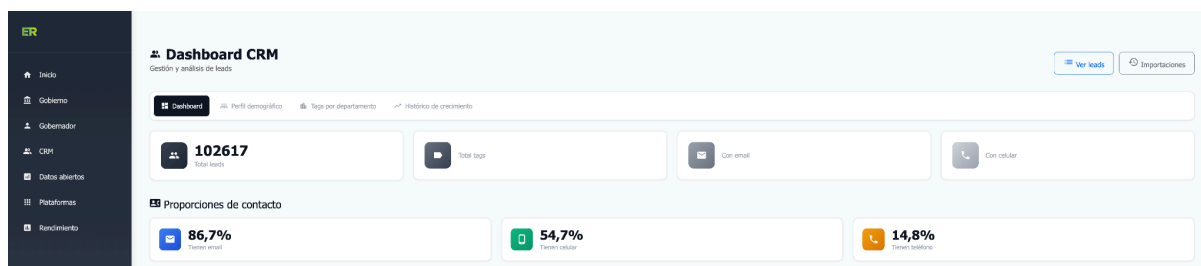
## Tarea 2: Integración de nuevas fuentes de datos en el sistema de monitoreo.

### 2.1. Incorporación de información de otros proyectos CFI y áreas de gobierno:

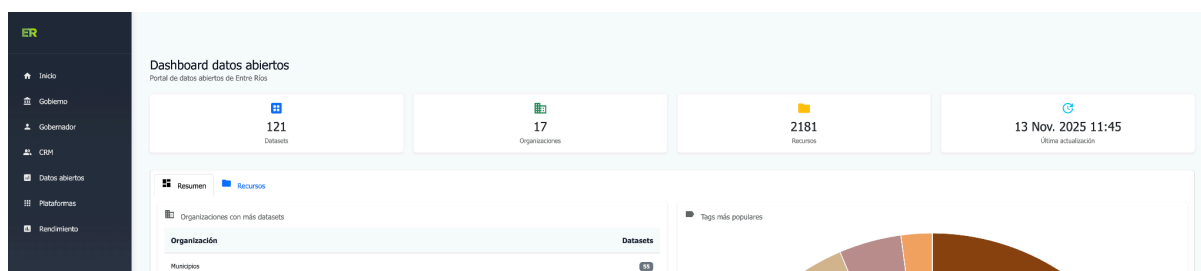
La incorporación de información externa se abordó como un eje de integración interproyecto: sumar a la lectura de comunicación datos que no provenían del circuito de métricas de redes o de tableros aislados, sino de sistemas y portales ya existentes en el ecosistema gubernamental. El objetivo fue reducir la fragmentación de consulta, integrando en la plataforma señales de interés ciudadano y de gestión que estaban disponibles en otros entornos.



En primer lugar, se integró información proveniente del CRM, entendido en el marco del proyecto como una fuente externa asociada a otro desarrollo del ecosistema CFI. Esta conexión permitió incorporar a la plataforma datos estructurados de leads y tags, y disponibilizar tanto vistas de consulta (listado con filtros) como vistas de trazabilidad de cargas (historial de importaciones) y analítica específica (perfil demográfico, tags por territorio e histórico de crecimiento). En términos operativos, esto habilita vincular el monitoreo de comunicación con dimensiones de segmentación y evolución de públicos que antes requerían consulta separada.



En segundo lugar, se incorporó información del portal provincial de datos abiertos, integrando su catálogo vía API CKAN. Dado que se trata de un sitio gubernamental externo a la plataforma, se desarrolló un conector para consultar datasets y recursos, normalizar metadatos y construir indicadores de seguimiento (totales, evolución temporal, formatos y checks de calidad básicos). Esto permite que el monitoreo no se limite a canales sociales, sino que incluya el estado de publicación y actualización de recursos informativos como parte del seguimiento.



Finalmente, para ordenar el acceso a plataformas gubernamentales complementarias, se implementó un módulo específico de “Plataformas” con hipervínculos directos a sitios externos de referencia. Esta pantalla funciona como punto de acceso unificado a recursos institucionales, evitando que el usuario tenga que salir a buscarlos en forma manual y manteniendo la navegación dentro del mismo entorno de trabajo.

Estos avances se evidencian en el módulo de CRM (tablero, leads, importaciones y analíticas), en el módulo de datos abiertos (inicio, exploración y detalle de datasets/recursos) y en la sección “Plataformas”, que centraliza enlaces de acceso directo a sitios gubernamentales.

## 2.2. Unificación y centralización de la información en un único tablero de control:

La unificación y centralización se abordó como un problema de arquitectura de datos y de producto: pasar de un esquema de consulta distribuida (herramientas separadas, tableros por fuente y consolidaciones manuales) a un sistema integrado donde las fuentes queden acopladas a un backend común, y donde los tableros compartan criterios homogéneos de cálculo, actualización y segmentación.

En términos de implementación, la plataforma se desarrolló sobre Django como capa de aplicación, con un enrutamiento por módulos que separa responsabilidades por dominio (métricas de redes, CRM, datos abiertos, métricas web y rendimiento de pauta) y una capa de autenticación para controlar el acceso a la información. La pantalla de inicio funciona como tablero general de síntesis, mientras que cada módulo expone tableros especializados con mayor granularidad. Este diseño permite sostener un punto de entrada único para lectura general y, al mismo tiempo, profundizar en cada fuente sin cambiar de herramienta.

En la capa de integración, se estandarizó el criterio de “fuente -> proceso de extracción/normalización -> capa de consulta -> visualización”. En el caso del CRM, la ingesta se implementó a partir de archivos CSV y un flujo de importación con registro de estado y cantidades procesadas, operado con transacciones para asegurar consistencia. La carga actualiza entidades normalizadas (leads y tags) mediante operaciones *idem-potentes* de actualización/creación, permitiendo reejecutar importaciones sin duplicar registros y manteniendo trazabilidad temporal. Cuando la fuente proviene de almacenamiento remoto, se incorporó un servicio de descarga desde S3 que copia los archivos al almacenamiento

de la aplicación, habilitando el mismo proceso de importación con independencia del origen del archivo.

Para datos abiertos, la integración se resolvió mediante consumo directo de la API CKAN del portal provincial. Se construyó un cliente con sesión HTTP, timeouts y reintentos, y se incorporó cache en puntos críticos de consulta para reducir latencia y desacoplar la navegación del usuario de la respuesta inmediata del servicio externo. Sobre esta base, se implementó una capa de servicio que agrega KPIs del catálogo y normaliza datasets/recursos para presentación consistente, incluyendo cálculos de totales, formatos predominantes, organizaciones activas, evolución temporal y chequeos de calidad básicos a nivel dataset (por ejemplo, presencia de licencia, descripción y completitud mínima de recursos).

Nombre del recurso	Dataset	Fecha creación	Organización	Tags	Formato
Septiembre 2025- Remuneraciones de autoridades su	Escala salarial de autoridades superiore	2025-11-13	Secretaría de Modernización	Escala salarial, Funcionarios, Sueldos	CSV
Octubre 2025- Remuneraciones de autoridades superi	Escala salarial de autoridades superiore	2025-11-13	Secretaría de Modernización	Escala salarial, Funcionarios, Sueldos	XLS
Octubre 2025- Remuneraciones de autoridades superi	Escala salarial de autoridades superiore	2025-11-13	Secretaría de Modernización	Escala salarial, Funcionarios, Sueldos	CSV
Septiembre 2025- Remuneraciones de autoridades sup	Escala salarial de autoridades superiore	2025-11-12	Secretaría de Modernización	Escala salarial, Funcionarios, Sueldos	XLS
Juntas Evaluadoras de Discapacidad de Entre Ríos	Juntas Evaluadoras de Discapacida en la	2025-11-12	Instituto Provincial de Discap	Discapacidad, Juntas evaluadoras	XLS
Juntas Evaluadoras de Discapacidad de Entre Ríos	Juntas Evaluadoras de Discapacida en la	2025-11-12	Instituto Provincial de Discap	Discapacidad, Juntas evaluadoras	CSV
Julio 2025- Remuneraciones de autoridades superior	Escala salarial de autoridades superiore	2025-09-15	Secretaría de Modernización	Escala salarial, Funcionarios, Sueldos	XLS
Julio 2025- Remuneraciones de autoridades superior	Escala salarial de autoridades superiore	2025-09-15	Secretaría de Modernización	Escala salarial, Funcionarios, Sueldos	CSV
Agosto 2025- Remuneraciones de autoridades superio	Escala salarial de autoridades superiore	2025-09-15	Secretaría de Modernización	Escala salarial, Funcionarios, Sueldos	CSV
Agosto 2025- Remuneraciones de autoridades superio	Escala salarial de autoridades superiore	2025-09-15	Secretaría de Modernización	Escala salarial, Funcionarios, Sueldos	XLS

En métricas de redes sociales, la unificación se materializa en tableros por canal (Facebook/Instagram) que comparten una estructura de consulta por cuenta y ventana temporal, y exponen tanto KPIs como series de evolución y desagregaciones (interacciones, crecimiento de audiencia, demografía, rendimiento de piezas). Se adoptó un patrón de “vista HTML inicial + endpoints de API internos” para que ciertos componentes puedan actualizarse por consultas asincrónicas, separando la obtención de datos del renderizado. Esto permite ampliar el set de visualizaciones sin replicar lógica entre pantallas.

En el caso de métricas web y rendimiento de pauta, la centralización se sostiene en el mismo criterio: el tablero de métricas web actúa como módulo específico para KPIs de tráfico, y el módulo de rendimiento permite registrar y consultar rendiciones por mes y año con agregaciones automáticas, asegurando que el tablero general pueda incluir una síntesis de estos indicadores sin consolidaciones manuales.

Desde el punto de vista de datos, la estrategia de centralización se apoya en un repositorio relacional único y en el uso del ORM como interfaz de acceso. Esto permite

mantener un modelo de entidades consistente, aplicar índices y criterios de ordenamiento donde corresponde, y construir agregaciones en backend para evitar cálculos pesados del lado cliente. La separación en servicios (por ejemplo, procesamiento CRM, integración S3, cliente CKAN y servicios de agregación) encapsula la lógica de integración y facilita la evolución de cada conector sin afectar la capa de presentación.

Estos avances se evidencian en el tablero de inicio (síntesis transversal), y en la navegación por módulos: el tablero de métricas de Facebook e Instagram centraliza KPIs, evolución y segmentaciones por cuenta; el módulo de CRM concentra importaciones, listados y analítica (perfil demográfico, tags por territorio e histórico); el módulo de datos abiertos habilita KPIs, exploración y detalle de datasets/recursos con controles de calidad; el módulo web expone KPIs del sitio; y el módulo de rendimiento integra carga y agregación de rendiciones. En conjunto, la plataforma consolida en un único entorno la consulta y gestión de indicadores previamente dispersos.

### **2.3. Optimización del procesamiento y almacenamiento de datos:**

La optimización del procesamiento y almacenamiento se trabajó con un enfoque de estabilidad y eficiencia, priorizando que la plataforma pueda actualizar y servir métricas sin depender de operaciones manuales y sin introducir complejidad innecesaria para el volumen de datos manejado. En este sentido, se adoptó una estrategia de repositorio relacional único y procesos de actualización reproducibles, alineados al uso operativo de los tableros.

Como backend de persistencia, se utiliza un motor relacional liviano (SQLite) integrado al core de la aplicación como fuente de verdad. La decisión de centralizar en un RDBMS único se complementa con el uso del ORM de Django como interfaz de acceso, lo que permite:

1. Mantener modelos explícitos para las entidades principales (por ejemplo, leads, tags, importaciones y rendiciones).
2. Evitar consultas SQL artesanales en la capa de aplicación.
3. Controlar índices y criterios de ordenamiento desde el modelo para mejorar tiempos de respuesta en consultas frecuentes.

En el módulo de CRM, la optimización se implementa en dos planos. Por un lado, el proceso de importación se ejecuta dentro de una transacción atómica, garantizando consistencia entre cargas de leads y tags y evitando estados parciales cuando ocurre un error. Por otro lado, se utiliza un patrón de actualización idempotente (actualización/creación por identificador original) que permite reejecutar importaciones sin duplicación y con trazabilidad temporal. El sistema registra el estado, los conteos procesados y los errores, permitiendo auditar corridas y observar su evolución en el historial de importaciones.

Para sostener performance en la navegación y la exploración de grandes listados, se incorporó paginación en vistas de consulta (por ejemplo, listado de leads e historial de importaciones), y se construyeron agregaciones en backend para alimentar tableros sin trasladar cálculos al frontend. En las vistas analíticas del CRM (perfil demográfico, tags por

departamento e histórico de crecimiento), los cálculos se apoyan en agregaciones y conteos sobre el repositorio relacional, lo que permite reutilizar los mismos datos normalizados en diferentes recortes.

En el módulo de datos abiertos, la optimización se resuelve principalmente en la capa de integración externa: se implementan reintentos con backoff ante fallos de red, control de timeouts y cache de resultados para búsquedas y consultas de detalle. Esto desacopla la experiencia de uso de la plataforma de la variabilidad de latencia de la API CKAN, y permite que operaciones comunes (exploración, filtros, detalle de dataset y previsualización) mantengan tiempos de respuesta consistentes. A su vez, la normalización de fechas, formatos y metadatos se realiza en una capa de servicio, evitando duplicación de lógica entre vistas.

En métricas de redes, la optimización se apoya en una separación entre renderizado inicial y endpoints internos para obtención de datos, permitiendo que componentes como KPIs, evolución y desagregaciones se actualicen sin recargar completamente la página. Esta arquitectura también permite incorporar nuevas visualizaciones manteniendo la misma estructura de consulta por cuenta y ventana temporal. En paralelo, el tablero de rendimiento de pauta consolida cálculos de acumulado y series mensuales en backend, con agregaciones directas sobre el modelo de rendiciones.

Finalmente, el diseño por servicios y módulos reduce el acoplamiento entre fuentes: los conectores (por ejemplo, CRM/almacenamiento remoto y CKAN) encapsulan su propia lógica de acceso, validación y normalización; las vistas consumen resultados ya estructurados; y los templates se enfocan en presentación. Estos criterios se evidencian en la estabilidad de navegación entre el tablero de inicio, los módulos de CRM (incluyendo históricos y segmentaciones), el módulo de datos abiertos (KPIs, exploración y detalle) y los tableros de redes, que operan sobre un repositorio de datos consistente y con procesos de actualización controlados.

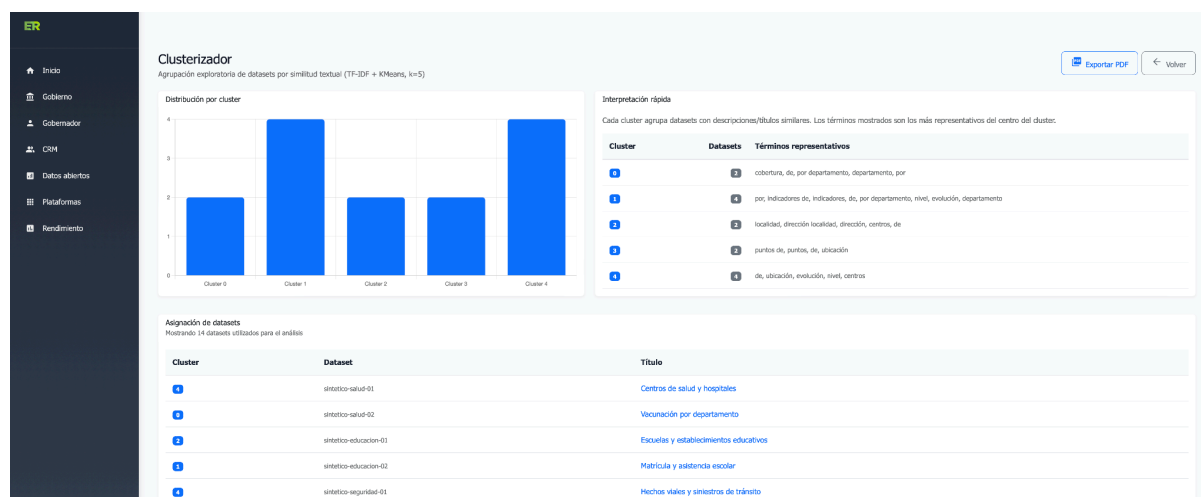
### **Tarea 3: Optimización de la plataforma de escucha activa y tendencias.**

#### **3.1. Nuevas capacidades avanzadas de Machine Learning:**

Para incorporar una capacidad real y visible de machine learning dentro de la plataforma, se implementó un componente de clusterización exploratoria aplicado al catálogo de datasets del portal CKAN. La lógica del desarrollo fue sumar una herramienta de lectura asistida que permita identificar grupos temáticos dentro del conjunto de publicaciones, sin necesidad de preclasificaciones manuales ni taxonomías previas.

La implementación se basó en un pipeline estándar y liviano: se construye un texto por dataset combinando título y descripción (metadatos title y notes), se vectoriza con TF-IDF y se calcula la agrupación mediante KMeans con  $k=5$ . El resultado no busca “predecir” ni reemplazar categorizaciones oficiales, sino generar una segmentación automática para exploración rápida: qué conjuntos de datasets tienden a parecerse entre sí por el lenguaje usado en su descripción.

En términos de evidencia dentro del sistema, la funcionalidad se observa en el módulo Datos abiertos, mediante un botón “Clusterizador” incorporado en el dashboard del módulo. La pantalla Datos abiertos → Clusterizador presenta una visualización de la distribución de datasets por cluster, una tabla de referencia con términos representativos por grupo (para interpretar el tema dominante de cada cluster) y el listado de datasets con su asignación para navegar al detalle cuando se requiere validación manual.



Adicionalmente, para asegurar que la capacidad sea demostrable en entornos donde el catálogo no responda o no haya metadatos suficientes para generar un vocabulario estable, el módulo contempla un modo de visualización con datos sintéticos, manteniendo el mismo pipeline de TF-IDF + KMeans. Esto permite sostener la evidencia visual del enfoque (gráfico y tablas) como parte de la entrega, sin alterar la lógica general ni el objetivo exploratorio de la tarea.

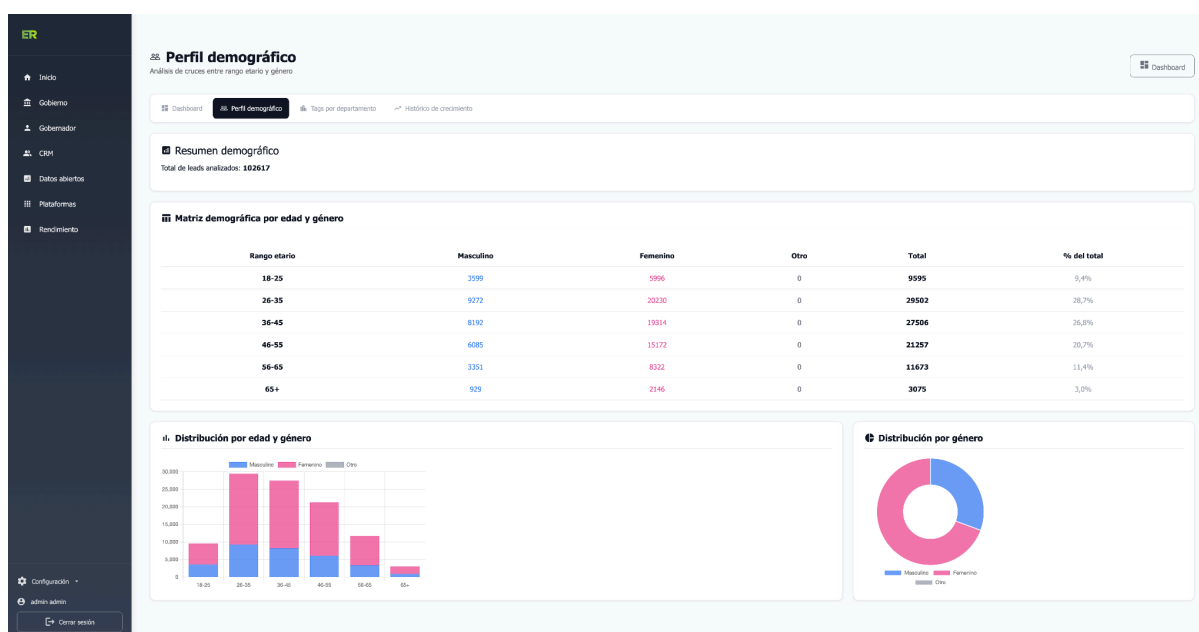
### 3.2. Ampliación de la cobertura de escucha activa:

La ampliación de la cobertura de escucha activa se abordó como una extensión del alcance de observación y no únicamente como un agregado de tableros. El objetivo fue pasar de una lectura concentrada en métricas tradicionales de redes a un esquema donde la “escucha” incorpore señales complementarias de interacción ciudadana y consumo de información, integrando fuentes que permiten interpretar comportamiento, interés y respuesta a acciones de comunicación.

En redes sociales, la cobertura se amplía desde un enfoque de KPIs básicos hacia una lectura más completa por cuenta y ventana temporal, incorporando series de evolución, distribución de interacciones, crecimiento de audiencia y segmentaciones de audiencia (demografía y ubicaciones). Este enfoque permite observar variaciones en el tiempo, detectar cambios de composición de audiencia y contextualizar resultados de piezas específicas dentro de una tendencia. En la plataforma, esto se evidencia en los módulos de

métricas de Facebook e Instagram, que permiten seleccionar cuentas y periodos, y acceder a desagregaciones y listados de rendimiento.

Como segundo eje, se incorporó al concepto de escucha la información proveniente del CRM, especialmente para analizar la respuesta a campañas y acciones de contacto masivo. La disponibilidad de leads y tags, junto con el registro de importaciones y el histórico de crecimiento, permite medir volumen, evolución y segmentación de públicos, y cruzar señales de interés (por ejemplo, tags asociados a campañas) con el comportamiento observado. En la plataforma, esto se evidencia en el módulo de CRM, tanto en el tablero principal como en las vistas analíticas (perfil demográfico, tags por territorio e histórico).

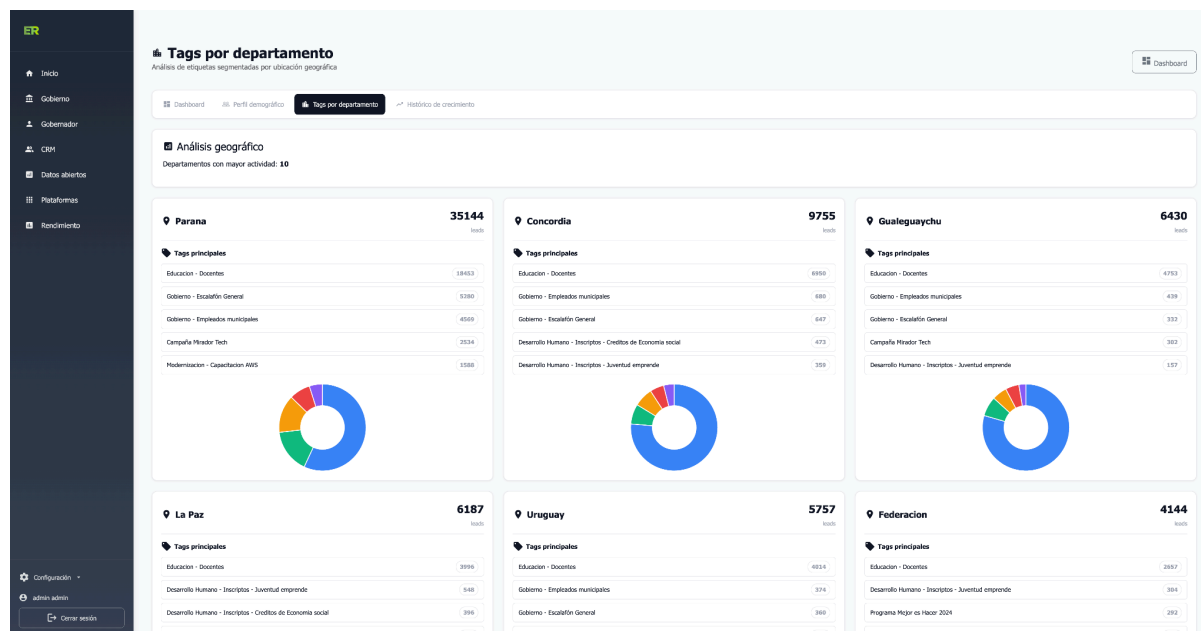


En paralelo, se incorporó la dimensión de datos abiertos como una señal adicional de “consumo de información” por parte de la ciudadanía y de cumplimiento institucional. Al integrar el catálogo CKAN, la plataforma puede monitorear KPIs del catálogo (datasets, recursos, organizaciones), evolución temporal y distribución por formatos/tags, además de visualizar el detalle por dataset y recurso. Esta información permite incluir en el esquema de escucha variables que no estaban presentes en el monitoreo tradicional de redes, aportando una lectura complementaria sobre qué recursos se publican, con qué calidad mínima y con qué frecuencia se actualizan. En la plataforma, esto se evidencia en el módulo de datos abiertos, en el tablero de KPIs, en el explorador con filtros y en las pantallas de detalle.

Finalmente, el criterio de cobertura se consolida en la pantalla de inicio, que integra una síntesis de indicadores por dominio y funciona como punto de entrada para detectar variaciones relevantes y derivar a los módulos específicos. De este modo, la escucha activa deja de ser un único tablero por canal y pasa a operar como una lectura integral, donde redes, CRM, datos abiertos, web y rendimiento pueden interpretarse en conjunto según el tipo de análisis requerido.

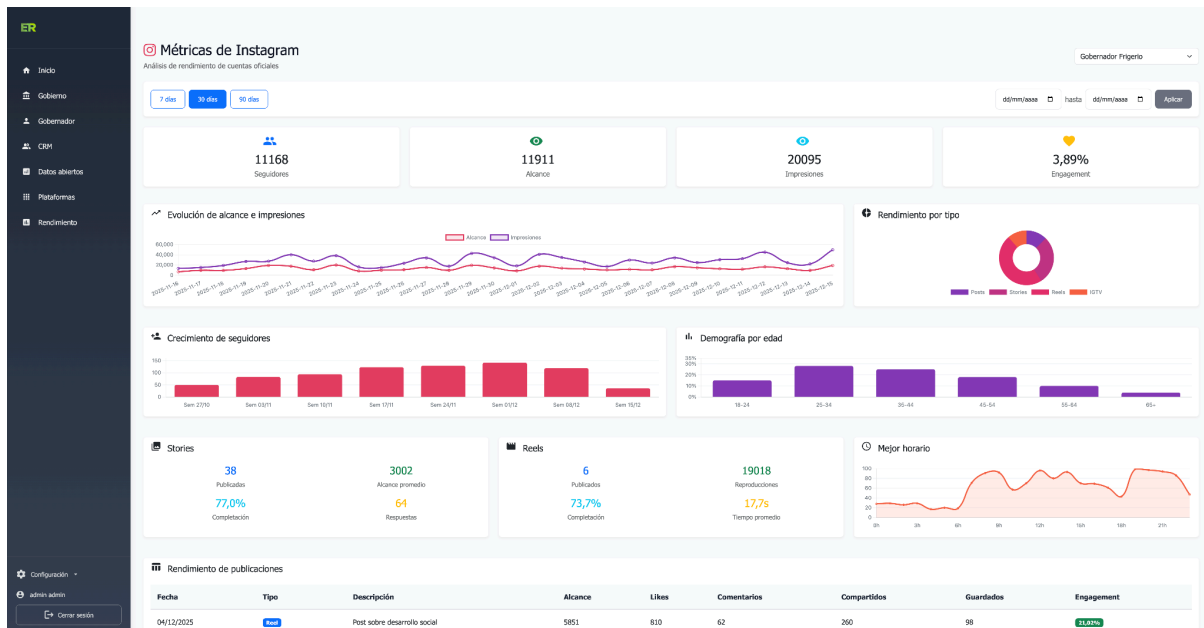
### 3.3. Mejora de la visualización de tendencias y segmentación de datos:

La mejora de visualización y segmentación se trabajó como una capa transversal a los módulos, con el objetivo de que las métricas integradas no se presenten únicamente como valores aislados sino como series interpretables y explorables. El foco estuvo puesto en dos aspectos: por un lado, habilitar lecturas temporales comparables (tendencias, variaciones y evolución) y, por otro, permitir cortes de análisis por dimensiones relevantes (cuentas, períodos, demografía, territorio, tags y categorías).



En métricas de redes, las pantallas de Facebook e Instagram se organizaron alrededor de un esquema de consulta parametrizable (selección de cuenta y ventana de días), y se estructuraron visualizaciones que combinan KPIs con series de evolución y desagregaciones. En términos técnicos, se implementó un patrón de renderizado que separa una carga inicial de la vista y endpoints internos para actualizar componentes puntuales (KPIs, evolución, rendimiento de publicaciones y demografía) sin recargar la página completa. Esto reduce el costo de refresco de datos y permite incorporar nuevos bloques de visualización manteniendo el mismo contrato de parámetros.

La segmentación en redes se implementa mediante desagregaciones de audiencia (edad/género, ubicaciones) y por tipo de contenido (por ejemplo, piezas destacadas), que habilitan análisis de cambios de composición y lecturas por comportamiento. La lógica de segmentación se mantiene del lado de los servicios de métricas, que entregan estructuras listas para ser consumidas por los componentes de visualización, evitando cálculos pesados en el frontend.



En CRM, las mejoras se enfocaron en habilitar exploración sobre listados y tableros analíticos sin perder consistencia. El listado de leads incorpora filtros por variables principales (provincia, departamento, sexo, rangos de edad y búsqueda) y paginación para sostener performance. Sobre la misma base, se construyeron vistas analíticas específicas para segmentación: perfil demográfico (cruces de rangos de edad y género), lectura territorial de tags por departamento, e histórico de crecimiento a partir de importaciones exitosas. Este diseño permite que una misma fuente normalizada soporte tableros de síntesis, exploración y análisis sin replicar estructuras.

En datos abiertos, la visualización se estructuró en tres niveles: un tablero de KPIs con métricas agregadas del catálogo, un explorador con filtros y ordenamiento, y pantallas de detalle por dataset/recurso. A nivel técnico, la capa de servicios normaliza formatos y fechas, construye agregados (formatos populares, organizaciones activas, evolución temporal, tags) y expone chequeos de calidad básicos en la vista de detalle. Para sostener tiempos de respuesta, se utiliza cache en consultas recurrentes y paginación/limitado en resultados listados.

Finalmente, la coherencia visual y operativa se consolida en el tablero de inicio, que actúa como síntesis de módulos y como punto de navegación hacia recortes más específicos. En conjunto, estas mejoras permiten pasar de tableros aislados a un esquema de lectura por tendencias y segmentación, donde los usuarios pueden alternar entre síntesis y detalle manteniendo criterios consistentes de filtros, periodos y definición de indicadores.

## **Tarea 4: Capacitación y transferencia de conocimiento.**

### **4.1. Expansión del programa de formación a nuevos equipos gubernamentales:**

La capacitación se abordó como un componente clave para asegurar adopción y continuidad del uso de la plataforma más allá del equipo inicial. En este marco, se amplió el programa de formación hacia nuevos equipos técnicos y analistas del área, así como equipos de áreas asociadas, con el objetivo de estandarizar el uso operativo de los tableros y sostener criterios comunes de lectura de métricas.

El proceso de transferencia se enfocó en el uso práctico de la plataforma: navegación por módulos, comprensión de KPIs y tendencias, aplicación de filtros y recortes temporales, e interpretación de desagregaciones por cuenta, territorio, audiencia y tags. A su vez, se trabajó sobre lineamientos de trabajo para ordenar la operación: criterios para definir periodos de análisis, rutinas de revisión, validación de lecturas en pantallas de detalle, y buenas prácticas para documentar hallazgos y compartir recortes consistentes entre equipos.

En términos de evidencia, esta transferencia se apoya en el funcionamiento integral del sistema: desde el tablero de inicio como síntesis transversal, hasta los módulos específicos (CRM, métricas por canal, datos abiertos, web y rendimiento) donde se ejercitan los criterios de lectura y segmentación. De este modo, la plataforma no sólo consolida datos y visualizaciones, sino que habilita un marco de trabajo común para que distintos equipos puedan operar con autonomía sobre los mismos indicadores.

### **4.2. Desarrollo de módulos avanzados de capacitación:**

Para sostener la expansión del programa de formación y evitar que el uso de la plataforma se reduzca a una navegación superficial por tableros, se desarrollaron módulos pedagógicos orientados a profundizar competencias analíticas y criterios de lectura. Estos módulos se estructuraron como unidades de trabajo aplicadas a pantallas reales del sistema, combinando explicación conceptual, ejercicios guiados y validación de resultados.

Entre los contenidos trabajados se incluyeron conceptos avanzados para operar con métricas de forma consistente: definición y alcance de KPIs (qué mide y qué no mide cada indicador), lectura de series temporales (tendencia, estacionalidad y variaciones puntuales), comparabilidad por periodo (ventanas temporales homogéneas y cortes equivalentes), y segmentación por dimensiones (cuenta, territorio, demografía y tags) para pasar de un valor agregado a una interpretación accionable.

Asimismo, se incorporaron lineamientos de calidad y trazabilidad para fortalecer la interpretación: criterios de completitud y normalización de datos, detección de inconsistencias entre fuentes, uso de pantallas de detalle para validar hallazgos (por ejemplo, navegación desde listados a detalle en CRM o desde exploración a detalle de

datasets/recursos en datos abiertos), y armado de evidencias reproducibles mediante recortes y salidas exportables cuando corresponde.

De este modo, la capacitación no se limitó a la adopción de la herramienta, sino que consolidó un marco metodológico para la lectura de métricas y la construcción de diagnósticos, alineando a equipos técnicos y analistas en un mismo lenguaje operativo.

## **Conclusiones**

El proyecto permitió consolidar un cambio de enfoque en el monitoreo de comunicación: de un esquema distribuido, apoyado en tableros y consultas por fuente, hacia una plataforma integral que centraliza fuentes, normaliza datos y expone métricas con criterios consistentes. Esta consolidación no se limita a la visualización, sino que incorpora procesos de integración y actualización que sostienen la operación cotidiana del monitoreo.

Desde el punto de vista técnico, la solución se apoya en una arquitectura modular sobre Django, con un repositorio relacional único y una capa de servicios que encapsula la lógica de integración (CRM, datos abiertos, métricas por canal y rendimiento). Esto permite mantener separadas las responsabilidades de extracción, normalización, consulta y presentación, y facilita la evolución de cada conector sin afectar el resto de la plataforma.

En términos de uso, la plataforma habilita lecturas comparables por periodo y recortes de segmentación que antes requerían consolidaciones manuales. La pantalla de inicio actúa como síntesis transversal y los módulos específicos permiten profundizar en cada dominio, combinando KPIs, evolución temporal y desagregaciones relevantes (por cuenta, audiencia, territorio, tags o calidad de publicación).

Finalmente, el sistema deja una base estable para nuevas iteraciones: incorporación de fuentes adicionales, ampliación del set de métricas y profundización de análisis, manteniendo un esquema de datos unificado y un circuito de monitoreo trazable. En ese sentido, el principal resultado del proyecto es disponer de una infraestructura operativa para tomar decisiones de comunicación con información consolidada y consultable en un único entorno.

## ANEXO I - DESCRIPCIÓN DE ARQUITECTURA

### 1. *Visión general*

La aplicación *er\_comunicacion* está desarrollada en Django, con un proyecto principal (*er\_comunicacion*) y una app central (*core*). Se utiliza el patrón MVT: las vistas orquestan el acceso a la base de datos vía ORM y renderizan templates HTML. Toda la información operativa se centraliza en una base relacional SQLite (*db.sqlite3*), que actúa como backend único para dashboards y reportes.

### 2. *Capa de presentación*

La capa de presentación se implementa con el motor de templates de Django, usando un template base y templates específicos en *templates/core*. Los recursos estáticos (CSS y JS) se organizan en *static/*, configurados mediante *STATIC\_URL*, *STATICFILES\_DIRS* y *STATIC\_ROOT*. Los gráficos y tableros consumen datos ya agregados en el backend y solo realizan lógica ligera de presentación en JavaScript. Crispy Forms con bootstrap5 se usa para formularios y pantallas de autenticación/administración.

### 3. *Capa de aplicación y servicios*

La lógica de aplicación reside en *core.views* y en servicios especializados bajo *core/services*. Las vistas protegidas con *login\_required* generan:

- Un dashboard general con métricas de leads y rendiciones de pauta.
- Tableros específicos para CRM, listados filtrables de leads e historial de importaciones.

La lógica de procesamiento de CRM se desacopla en la clase *CRMDataProcessor* (*crm\_processor.py*), que implementa:

- Métodos de orquestación de importaciones (*procesar\_importacion\_completa*, *procesar\_importacion\_local*).
- Procesos de lectura CSV con *pandas*, normalización de campos y escritura idempotente mediante *Lead.objects.update\_or\_create*.
- Cálculo de resúmenes agregados para alimentar dashboards (totales, distribución por provincia, sexo, tags, contacto).

La integración con S3 se abstrae en *S3CRMService* (*s3\_service.py*), que:

- Inicializa un cliente *boto3* a partir de variables de entorno.
- Lista y descarga archivos CSV de CRM (*leads.csv*, *leads\_tags.csv*) al storage de Django.
- Expone utilidades para limpiar archivos temporales luego de las importaciones.

### 4. *Modelo de datos*

El backend de datos se define exclusivamente a través del ORM en *core/models.py*:

- Usuario: modelo custom (AUTH\_USER\_MODEL) basado en AbstractUser, con campos de email único, nombre y apellido, utilizado tanto para autenticación como para trazabilidad (por ejemplo, en RendicionPauta.creado\_por).
- Lead: entidad central de CRM con sexo, edad, flags de contacto y localización. Incluye timestamps de creación/actualización/importación e índices por provincia/departamento, sexo/edad y fecha\_importacion.
- LeadTag: relación etiquetada entre leads y tags de texto, con unicidad (lead, tag) e índices por tag y fecha\_importacion, optimizada para análisis de segmentación y ranking de etiquetas.
- ImportacionCRM: registro histórico de procesos de importación con timestamps, estado, contadores y campo de errores; incluye una propiedad duracion para análisis de performance.
- RendicionPauta: registro de montos por mes/año y usuario creador, con choices de meses, ordenamiento descendente por periodo y método auxiliar para obtener años disponibles.

## *5. Integración, seguridad y despliegue*

La configuración de settings.py define middleware estándar de seguridad y sesiones, idioma es-ar, zona horaria America/Argentina/Buenos\_Aires, URLs de login/logout y el uso del modelo de usuario personalizado. En modo DEBUG, archivos estáticos y de medios se sirven vía Django mediante static(), mientras que en producción se espera un servidor web o entorno de contenedores que exponga STATIC\_ROOT y MEDIA\_ROOT.

En conjunto, la arquitectura combina:

- Backend relacional único (SQLite + ORM Django).
- Servicios dedicados para ingesta (pandas + CRMDataProcessor) e integración con S3 (S3CRMSERVICE).
- Vistas delgadas que consumen estos servicios y exponen dashboards y listados filtrables sobre datos ya normalizados

## **ANEXO II - DESCRIPCIÓN DE FUENTES**

### **1. CRM y gestión de contactos**

Fuente: Sistema CRM institucional (exportaciones CSV y/o integración vía S3).

Tipo de datos: registros de leads con atributos sociodemográficos y de contacto.

Estructura principal:

- Identificador original del lead (id\_original).
- Sexo, edad.
- Disponibilidad de canales de contacto: email, celular, teléfono (flags 0/1).
- Localización: provincia, departamento, localidad.
- Tags o etiquetas asociadas (LeadTag), que representan temas, campañas, canales u orígenes.

Modo de integración:

- Ingesta periódica de archivos leads.csv y leads\_tags.csv.
- Procesamiento mediante pandas y escritura a través del ORM de Django sobre los modelos Lead y LeadTag.
- Registro de cada corrida en ImportacionCRM con timestamps, estado y cantidad de registros procesados.

Uso analítico:

- Cálculo de totales de leads, distribución por provincia y departamento.
- Distribución por sexo y rangos de edad.
- Proporción de leads con email, celular y teléfono disponible.
- Ranking de tags más frecuentes y segmentación por etiquetas temáticas.

### **2. Métricas de redes sociales (Gobierno y Gobernador)**

Fuente: Plataformas de redes sociales oficiales (cuentas de Gobierno y Gobernador).

Tipo de datos: indicadores agregados de performance por periodo.

Estructura principal (a nivel de tablero):

- Facebook: alcance total, impresiones, otras métricas agregadas por página.
- Instagram: seguidores, alcance por periodo y métricas de interacción.

Modo de integración:

- Centralización de los valores agregados en el backend, asociados a cada módulo de redes.
- Consumo desde vistas específicas de dashboards de Gobierno y Gobernador, que presentan la evolución temporal y comparaciones entre cuentas.

Uso analítico:

- Monitoreo de alcance e impresiones por página y periodo.
- Comparación entre las cuentas de Gobierno y Gobernador.
- Lectura unificada de la performance digital junto con otros canales (web, CRM, pauta).

### **3. Métricas web del sitio del Gobierno**

Fuente: Herramienta de analítica web conectada al sitio oficial del Gobierno de Entre Ríos.

Tipo de datos: métricas de tráfico y comportamiento de usuarios.

Estructura principal:

- Visitas totales y usuarios únicos.
- Páginas vistas.
- Distribución por dispositivo y/o canal de tráfico (según disponibilidad de la fuente).

Modo de integración:

- Consolidación de métricas agregadas en el backend como módulo de “web” dentro del dashboard general.
- Exposición en tableros que permiten ver la evolución temporal y cruces con otras dimensiones de análisis.

Uso analítico:

- Evaluación del uso del sitio y consumo de contenidos.
- Cruce con campañas y acciones en redes sociales o CRM.
- Identificación de picos de tráfico asociados a eventos o acciones de comunicación.

### **4. Datos abiertos y consumo de recursos de información**

Fuente: Portal de datos abiertos (por ejemplo, instancia CKAN-Andino del Gobierno).

Tipo de datos: metadatos y métricas de uso de datasets publicados.

Estructura principal:

- Identificación de datasets y recursos.
- Métricas de consulta/descarga y ranking de datasets más accedidos (según datos expuestos por la API del portal).

Modo de integración:

- Ingesta vía APIs del portal de datos abiertos y normalización en modelos relacionales.
- Incorporación de los indicadores al módulo de datos abiertos en la plataforma, permitiendo visualizar el uso relativo de los recursos.

Uso analítico:

- Lectura del impacto de las políticas de apertura de datos.
- Cruce de consumo de datasets con campañas y acciones de comunicación.
- Identificación de temas de interés ciudadano a partir de los recursos más consultados.

### **5. Rendición de pauta y registro de inversión**

Fuente: Sistema de registro interno de rendición de pauta publicitaria.

Tipo de datos: montos rendidos por mes y año.

Estructura principal (modelo RendicionPauta):

- Monto (decimal).
- Mes (choice 1–12).

- Año.
- Usuario creador (referencia a Usuario).
- Fecha de creación.

Modo de integración:

- Carga y gestión desde la propia aplicación Django, persistida en la base relacional.
- Exposición de series temporales de montos rendidos y agregados por periodo.

Uso analítico:

- Seguimiento de la inversión en pauta por mes y año.
- Cruce con métricas de alcance e impacto de redes, CRM y web.
- Apoyo a análisis de eficiencia y planificación presupuestaria en comunicación.

## **6. Síntesis de integración**

Todas las fuentes anteriores se consolidan en un único backend relacional (SQLite) a través de pipelines específicos (CRM + S3, datos abiertos, métricas agregadas de redes y web, rendición de pauta). Sobre esta base unificada se construyen los dashboards y vistas de la plataforma er\_comunicacion, que permiten analizar de forma integrada:

- Performance de comunicación en redes y web.
- Comportamiento y segmentación de públicos (CRM + tags).
- Consumo de recursos de información (datos abiertos).
- Relación entre inversión en pauta y resultados observados en los distintos canales.

Aunque algunas fuentes utilicen datos sintéticos o de prueba en el entorno actual, la estructura de modelos, pipelines y dashboards está preparada para operar con datos productivos sin cambios en la arquitectura.

### **ANEXO III - CONCEPTOS FUNDAMENTALES DE LOS PROCESOS DE TRANSFERENCIA**

- Marco de medición en comunicación digital: objetivos de comunicación, objetivos de campaña y su traducción operativa en indicadores.
- Diferencia entre indicadores de exposición, interacción, conversión y retención: lectura por etapa del embudo y riesgos de interpretar KPIs fuera de contexto.
- Definición de KPI: variable, unidad, ventana temporal, población, fuente de datos y regla de cálculo.
- Métricas primarias vs métricas derivadas: consistencia de definiciones, trazabilidad de fórmulas y control de comparabilidad.
- Normalización de métricas entre canales: equivalencias, límites de comparabilidad y criterios de homologación por ventana temporal.
- Series temporales aplicadas a métricas de comunicación: tendencia, estacionalidad, ruido, outliers y eventos.
- Ventanas de análisis: “últimos N días” vs recortes calendario; implicancias en comparabilidad interperiodo.
- Lectura de variación: variación absoluta, variación relativa, tasa de crecimiento y normalización por base.
- Interpretación de picos: identificación de drivers (campaña, pauta, viralidad, publicación específica, evento externo) y validación con evidencias.
- Sesgos de interpretación: sesgo de supervivencia, sesgo por cambios de base, cambios de algoritmo/plataforma, estacionalidad no controlada.
- Segmentación analítica: definición de dimensiones de corte (cuenta, territorio, audiencia, tags, tipo de contenido) y selección de granularidad.
- Cohortes y comparación por segmentos: lectura de desempeño relativo entre segmentos, estabilidad de muestra y control de tamaño mínimo.
- Métricas de audiencia: composición por edad/género, distribución geográfica, lectura de cambios de mix y su impacto en performance.
- Métricas de contenido: tipologías de piezas, distribución de formatos, rendimiento por formato y análisis de saturación.
- Métricas de alcance e impresiones: definiciones, solapamientos y lectura de cobertura vs frecuencia.
- Interacciones: desagregación por tipo (reacciones, comentarios, compartidos) y lectura cualitativa complementaria.

- Ratios de eficiencia: engagement rate, interacción por alcance, interacción por impresiones; criterios de uso y limitaciones.
- Conversión y respuesta: relación entre métricas de redes y señales internas (CRM/leads); criterios de atribución exploratoria.
- Gobernanza de indicadores: glosario, versionado de definiciones, responsables de cálculo y auditoría de cambios.
- Datos abiertos como dimensión de escucha: lectura de publicación y actualización como señal institucional.
- Catálogo CKAN: conceptos de dataset, recurso, organización, licencia, formato y metadatos.
- KPIs de catálogo: total de datasets, total de recursos, organizaciones activas, distribución por formatos.
- Evolución temporal en datos abiertos: lectura de "metadata\_modified" y límites de interpretación (metadato vs actualización real del recurso).
- Calidad en datos abiertos: completitud de metadatos (título, descripción, licencia), consistencia de formatos y presencia de recursos.
- Chequeos de calidad básicos: detección de recursos sin formato, links rotos (cuando aplica), descripciones vacías.
- Exploración con filtros: búsqueda por texto, filtros por organización y ordenamientos; criterios para recortes comparables.
- Lectura por detalle: interpretación de ficha de dataset, recursos disponibles y previsualización.
- Introducción a analítica exploratoria aplicada: hipótesis, validación rápida y límites de generalización.
- Clusterización exploratoria (caso datos abiertos): similitud textual, TF-IDF, KMeans y lectura de clusters como agrupación temática.
- Interpretación de términos representativos: lectura del "centro" del cluster, ruido de vocabulario y necesidad de validación manual.
- Buenas prácticas en ML exploratorio: reproducibilidad (semilla), selección de k, sensibilidad a datos, y uso como herramienta de apoyo.