# PROPUESTA TÉCNICA INTEGRAL

## Sistema de Centralización y Visualización de Datos EMTP

**Secretaría Ejecutiva de Educación Media Técnico Profesional (SEEMTP)**  
**Ministerio de Educación de Chile**

## 📋 RESUMEN EJECUTIVO

La presente propuesta integra dos componentes esenciales para la modernización de la gestión de datos de la SEEMTP:

1. **FASE I: Centralización y Estandarización de Datos** - Infraestructura base para consolidar información dispersa en SharePoint y múltiples fuentes
2. **FASE II: Plataforma de Visualización Interactiva** - Sistema web avanzado (Python/Dash) para análisis dinámico y toma de decisiones basada en evidencia

**Beneficio Integral:** Un ecosistema completo que transforma datos dispersos y no estructurados en información accionable mediante visualizaciones interactivas, alertas automáticas y análisis multidimensional.

## 1. CONTEXTO Y PROBLEMÁTICA

### 1.1 Situación Actual

La SEEMTP enfrenta dos desafíos críticos e interrelacionados:

#### **Problema A: Dispersión de Datos de Gestión de Proyectos (Datos Internos)**

* **Ubicación:** SharePoint > Carpeta General SEEMTP > 07\_Equipo Gestión
* **Estructura actual:**
  + 001\_EMTP
  + 002\_EQUIPAMIENTO\_EMTP
  + 003\_EQUIPAMIENTO\_SLEP
  + 004\_APOYO\_SLEP

**Impactos identificados:** - ❌ Planillas Excel sin estandarización entre años (mismo proyecto, campos distintos) - ❌ Datos incompletos, duplicados e inconsistentes - ❌ Falta de alertas sobre vencimientos de rendiciones - ❌ Imposibilidad de consolidar reportes comparativos - ❌ Pérdida de trazabilidad en gestión de proyectos

**Responsable de ordenar:** Equipo de Gestión SEEMTP

#### **Problema B: Información Educativa Desagregada (Datos SIGE MINEDUC)**

Datos disponibles en bases de datos institucionales (SIGE, SIES) que requieren integración:

1. Caracterización de estudiantes y establecimientos EMTP
2. Docentes del sistema EMTP
3. Egresados/as y Titulados/as
4. Egresados/as EMTP matriculados en ESUP

**Requisito:** Coordinación con TI MINEDUC para acceso a bases de datos internas y actualización automática.

**Consecuencias actuales:** - 🚫 No existen roles claros para gestión de datos internos - 🚫 Información educativa no integrada dificulta análisis estratégico - 🚫 Procesos manuales demoran toma de decisiones

### 1.2 Necesidad de una Solución Integral

Ambos problemas requieren una solución sistémica que: - **Unifique** datos dispersos en un repositorio centralizado - **Estandarice** formatos y procesos de carga - **Automatice** limpieza, validación y alertas - **Visualice** información de forma interactiva para distintos perfiles

## 2. OBJETIVOS

### 2.1 Objetivo General

Implementar un **sistema integral de centralización y visualización de datos** que consolide información de la SEEMTP, automatice procesos de gestión y facilite la toma de decisiones basada en evidencia mediante dashboards interactivos.

### 2.2 Objetivos Específicos

#### **FASE I: Centralización**

1. Estandarizar estructura de datos de proyectos y rendiciones en SharePoint
2. Automatizar procesos ETL (Extract, Transform, Load) para limpieza y carga
3. Implementar sistema de alertas automáticas para vencimientos
4. Integrar datos de múltiples fuentes educativas en modelo unificado

#### **FASE II: Visualización**

1. Desarrollar plataforma web interactiva con dashboards dinámicos
2. Configurar accesos y permisos por rol de usuario
3. Habilitar exportación de reportes y análisis ad-hoc
4. Capacitar equipos SEEMTP en uso del sistema

## 3. ALCANCE DEL PROYECTO

### 3.1 Módulos de Datos a Centralizar

| Módulo | Fuente | Tipo de Dato | Responsable |
| --- | --- | --- | --- |
| **Matrícula EMTP** | SIGE MINEDUC (BD institucional) | Estudiantes, evolución anual, género, región, especialidad | TI MINEDUC (conexión automática) |
| **Titulación** | SIGE MINEDUC | Egresados, titulados, tasas por cohorte | TI MINEDUC (conexión automática) |
| **EMTP en ESUP** | SIES, cruce MINEDUC | Transición a educación superior, carreras elegidas | TI MINEDUC (conexión automática) |
| **Establecimientos** | Directorio Oficial MINEDUC | RBD, dependencia, SLEP, especialidades impartidas | TI MINEDUC (conexión automática) |
| **Docentes** | Registros MINEDUC, CPEIP | Perfil, contrato, formación, estabilidad laboral | TI MINEDUC (conexión automática) |
| **Proyectos SEEMTP** | SharePoint (07\_Equipo Gestión) **[INTERNO]** | Asignación recursos, rendiciones, estado ejecución | **Equipo Gestión SEEMTP (ordenar datos)** |

### 3.2 Funcionalidades del Sistema

#### **Centralización (Backend)**

* ✅ Repositorio único en SharePoint estructurado o BD institucional
* ✅ Procesos ETL automáticos con Python/Power Automate
* ✅ Validación de integridad y detección de duplicados
* ✅ Alertas automáticas por correo y Teams (rendiciones, vencimientos)
* ✅ Auditoría de cambios y trazabilidad completa

#### **Visualización (Frontend)**

* ✅ Dashboards interactivos con filtros dinámicos
* ✅ Análisis multidimensional (región, año, especialidad, género, dependencia)
* ✅ Gráficos comparativos y evolutivos (tendencias, benchmarking)
* ✅ Exportación a Excel/PDF de reportes personalizados
* ✅ Acceso web responsive (desktop, tablet, móvil)
* ✅ Autenticación segura con perfiles diferenciados

## 4. METODOLOGÍA DE IMPLEMENTACIÓN

### 4.1 Arquitectura Tecnológica Propuesta

┌─────────────────────────────────────────────────────────────────┐  
│ CAPA DE FUENTES DE DATOS │  
├─────────────────────────────────────────────────────────────────┤  
│ SIGE (BD MINEDUC) │ SIES │ SharePoint (Proyectos SEEMTP) │  
└────────────┬────────────────────────────────────────────────────┘  
 │  
 ▼  
┌─────────────────────────────────────────────────────────────────┐  
│ CAPA DE INTEGRACIÓN Y TRANSFORMACIÓN │  
├─────────────────────────────────────────────────────────────────┤  
│ • Conexión automática a SIGE (coordinación TI MINEDUC) │  
│ • ETL para datos de proyectos (Python, ordenados por Gestión) │  
│ • Limpieza y validación de datos │  
│ • Estandarización de formatos │  
│ • Detección de duplicados y anomalías │  
└────────────┬────────────────────────────────────────────────────┘  
 │  
 ▼  
┌─────────────────────────────────────────────────────────────────┐  
│ CAPA DE ALMACENAMIENTO CENTRAL │  
├─────────────────────────────────────────────────────────────────┤  
│ • Conexión directa a BD SIGE (datos educativos) │  
│ • Base de Datos para Proyectos SEEMTP (datos internos) │  
│ • Modelo de datos normalizado y documentado │  
└────────────┬────────────────────────────────────────────────────┘  
 │  
 ▼  
┌─────────────────────────────────────────────────────────────────┐  
│ CAPA DE APLICACIÓN │  
├─────────────────────────────────────────────────────────────────┤  
│ • API REST (Python/Flask/Dash) │  
│ • Validaciones y reglas de transformación de datos │  
│ • Sistema de alertas (correo, Teams) para proyectos │  
│ • Control de acceso y auditoría │  
└────────────┬────────────────────────────────────────────────────┘  
 │  
 ▼  
┌─────────────────────────────────────────────────────────────────┐  
│ CAPA DE VISUALIZACIÓN (FRONTEND) │  
├─────────────────────────────────────────────────────────────────┤  
│ • Aplicación Web Dash/Python (Dashboard interactivo) │  
│ • Gráficos Plotly (interactividad avanzada) │  
│ • Filtros dinámicos y exportación │  
│ • Autenticación y perfiles de usuario │  
└─────────────────────────────────────────────────────────────────┘

### 4.2 Stack Tecnológico

| Componente | Tecnología Propuesta | Justificación |
| --- | --- | --- |
| **Backend** | Python 3.11+ | Ecosistema robusto para análisis de datos, integración nativa con Pandas/NumPy |
| **Framework Web** | Dash 2.x (Plotly) | Especializado en visualización interactiva, desarrollo ágil de dashboards |
| **Base de Datos** | SQL Server (MINEDUC) o PostgreSQL | Compatibilidad con infraestructura existente, soporte institucional |
| **ETL** | Python (Pandas, SQLAlchemy) + Power Automate | Automatización robusta, integración con SharePoint |
| **Visualización** | Plotly.js | Gráficos interactivos de alta calidad, exportables |
| **Autenticación** | JWT + bcrypt | Seguridad estándar industria, compatible con SSO MINEDUC |
| **Deployment** | Docker + Gunicorn | Portabilidad, escalabilidad, facilita deployment en servidores institucionales |
| **Monitoreo** | Loguru | Trazabilidad completa de operaciones y errores |

**Ventaja vs Power BI:** Mayor control sobre procesamiento de datos, personalización ilimitada, no requiere licencias adicionales de Microsoft, código abierto y auditable.

## 5. FASES DE IMPLEMENTACIÓN

### **FASE I: INTEGRACIÓN Y CENTRALIZACIÓN DE DATOS** (4 semanas)

#### **Semana 1: Coordinación y Levantamiento**

**Responsable:** Equipo Políticas + Desarrollador Externo

**Actividades:** 1. **Coordinación con TI MINEDUC:** Solicitar accesos a BD SIGE (lectura) y documentación técnica 2. **Levantamiento datos internos:** Inventariar archivos Excel de proyectos en SharePoint (carpetas 001-004) 3. **Mapeo de campos:** Identificar variables comunes entre SIGE y datos internos 4. **Definir prioridades:** Determinar módulos críticos para implementación inicial

**Entregables:** - 📄 Solicitud formal de accesos a TI MINEDUC - 📄 Inventario de datos de proyectos (Excel) - 📄 Mapeo preliminar de campos SIGE → modelo aplicación

#### **Semana 2: Estandarización Datos Internos (Proyectos)**

**Responsable:** Equipo Gestión + Desarrollador Externo

**Actividades:** 1. **Estandarizar Excel de proyectos:** Equipo Gestión unifica campos entre años 2. **Definir diccionario de datos:** Campos obligatorios, formatos, validaciones 3. **Diseñar modelo para proyectos:** Estructura de tabla/BD para datos internos 4. **Implementar validaciones:** Scripts Python para detectar inconsistencias

**Entregables:** - 📄 Diccionario de datos de proyectos (versión 1.0) - � Excel estandarizados (2-3 años piloto) - � Scripts de validación y limpieza

#### **Semana 3-4: Conexión SIGE y Carga Datos**

**Responsable:** Desarrollador Externo + TI MINEDUC

**Actividades:** 1. **Configurar conexión a SIGE:** Credenciales, testing, documentación 2. **Desarrollar consultas SQL:** Extraer datos de matrícula, titulación, docentes, establecimientos 3. **Cargar datos de proyectos:** ETL para Excel → BD/estructura centralizada 4. **Testing integración:** Validar calidad y completitud de datos 5. **Sistema de alertas básico:** Configurar alertas de vencimientos (correo)

**Entregables:** - 🗄️ Conexión funcional a BD SIGE - 💻 Scripts ETL para datos de proyectos - 🔔 Sistema de alertas configurado (fase piloto) - 📄 Documentación técnica de conexiones

### **FASE II: DESARROLLO DE VISUALIZACIÓN** (4 semanas)

#### **Semana 5-6: Módulos Base y Dashboards Principales**

**Responsable:** Desarrollador Externo

**Actividades:** 1. Configurar aplicación Dash con autenticación 2. Desarrollar dashboards prioritarios: - **Matrícula EMTP** (datos SIGE) - **Proyectos SEEMTP** (datos internos con alertas visuales) - **Establecimientos y Docentes** (datos SIGE) 3. Implementar filtros dinámicos (región, año, especialidad) 4. Sistema de exportación básico (Excel)

**Entregables:** - 💻 Aplicación web funcional con 3 dashboards principales - 🔐 Sistema de autenticación por roles - 📊 Gráficos interactivos con Plotly

#### **Semana 7: Dashboards Complementarios**

**Responsable:** Desarrollador Externo

**Actividades:** 1. Dashboard Titulación y Egresados en ESUP (datos SIGE) 2. Refinamiento de filtros y navegación 3. Optimización de performance 4. Testing con usuarios piloto

**Entregables:** - 📊 5 módulos de visualización completos - 🔍 Filtros cruzados funcionales - 📄 Reporte de testing con usuarios

#### **Semana 8: Capacitación y Puesta en Marcha**

**Responsable:** Desarrollador Externo + Coordinación SEEMTP

**Actividades:** 1. Capacitación a equipos (gestión, políticas, usuarios finales) 2. Ajustes finales según feedback 3. Documentación de usuario y técnica 4. Deployment en servidor interno o cloud

**Entregables:** - 📄 Manual de usuario ilustrado - 📄 Manual técnico de administración - 🚀 Sistema en producción - 🎥 Video tutorial básico (opcional)

## 6. PERFIL DEL EQUIPO REQUERIDO

### 6.1 Desarrollador/a Externo/a (Contratación)

**Perfil técnico:** - ✅ Experiencia demostrable en Python (Pandas, Dash, Plotly) - ✅ Conocimiento de bases de datos relacionales (SQL Server/PostgreSQL) - ✅ Experiencia en desarrollo de aplicaciones web interactivas - ✅ Familiaridad con integración de datos (ETL, APIs, SharePoint) - ✅ Capacidad de documentación técnica clara

**Deseable:** - Experiencia en sector educativo o análisis de datos educativos - Conocimiento de visualización de datos estadísticos - Experiencia con deployment en servidores internos o cloud

**Dedicación:** 8 semanas tiempo completo

### 6.2 Equipo Interno SEEMTP

| Rol | Responsabilidad | Dedicación |
| --- | --- | --- |
| **Coordinador/a del Proyecto** | Supervisión general, coordinación con TI MINEDUC, aprobaciones | 10% tiempo |
| **Equipo Políticas** | Definición indicadores, validación cálculos, levantamiento requisitos | 20% tiempo |
| **Equipo Gestión** | Estandarización datos proyectos, validación información, testing | 20% tiempo |

**Coordinación externa:** TI MINEDUC (conexión a bases SIGE, sin costo adicional).

## 7. CRONOGRAMA Y PRODUCTOS ESPERADOS

### 7.1 Carta Gantt Resumida (8 semanas)

┌──────────────────────────────────────────────────────────────┐ │ FASE │ S1 │ S2 │ S3 │ S4 │ S5 │ S6 │ S7 │ S8 │ ├──────────────────────────────────────────────────────────────┤ │ FASE I: Integración Datos│ ██ │ ██ │ ██ │ ██ │ │ │ │ │ │ - Coordinación TI │ ██ │ │ │ │ │ │ │ │ │ - Datos internos │ │ ██ │ │ │ │ │ │ │ │ - Conexión SIGE + ETL │ │ │ ██ │ ██ │ │ │ │ │ │ │ │ │ │ │ │ │ │ │ │ FASE II: Visualización │ │ │ │ ██ │ ██ │ ██ │ ██ │ ██ │ │ - Dashboards base │ │ │ │ ██ │ ██ │ │ │ │ │ - Dashboards completos │ │ │ │ │ │ ██ │ │ │ │ - Capacitación │ │ │ │ │ │ │ ██ │ ██ │ └──────────────────────────────────────────────────────────────┘ ```

### 7.2 Entregables Clave

#### **FASE I - Integración Datos (4 semanas)**

1. 📄 Solicitud accesos TI MINEDUC
2. 📄 Diccionario de datos de proyectos
3. 💻 Scripts ETL para datos internos
4. 🗄️ Conexión funcional a BD SIGE
5. 🔔 Sistema de alertas básico
6. 📄 Documentación técnica de conexiones

#### **FASE II - Visualización (4 semanas)**

1. 💻 Aplicación web Dash completa
2. 📊 5 módulos de dashboards (matrícula, titulación, ESUP, establecimientos, docentes, proyectos)
3. � Sistema de autenticación
4. 📄 Manuales de usuario y técnico
5. 🚀 Sistema en producción

## 8. RIESGOS Y MITIGACIONES

| Riesgo | Probabilidad | Impacto | Mitigación |
| --- | --- | --- | --- |
| Calidad de datos históricos insuficiente | Media | Alto | Priorizar años con mejor calidad, proceso iterativo de limpieza |
| Resistencia al cambio de usuarios | Media | Medio | Capacitación intensiva, involucrar usuarios desde diseño |
| Integración compleja con sistemas MINEDUC | Alta | Alto | Coordinación temprana con TI, plan B con SharePoint estructurado |
| Cambios en estructura de datos de fuentes | Media | Medio | Capa de abstracción en ETL, monitoreo de cambios, alertas |
| Delays en accesos a BD institucionales | Alta | Alto | Iniciar trámites en paralelo, usar datos simulados para desarrollo |

## 9. SOSTENIBILIDAD DEL SISTEMA

### 9.1 Gobernanza Post-Implementación

**Roles definidos:**

| Rol | Responsabilidad | Frecuencia |
| --- | --- | --- |
| **Administrador de Datos** | Validar carga de datos nuevos, resolver inconsistencias | Semanal |
| **Administrador del Sistema** | Mantenimiento técnico, actualización de software | Mensual |
| **Gestor de Usuarios** | Gestión de accesos, permisos, capacitación nuevos usuarios | Según demanda |
| **Revisor de Alertas** | Seguimiento de alertas de vencimientos, escalamiento | Diario |

### 9.2 Actualización de Datos

**Flujo automatizado:**

1. **Fuentes SIGE MINEDUC:** Conexión directa (coordinada con TI)
2. **Proyectos SEEMTP (SharePoint):** ETL automático detecta nuevos archivos
3. **Datos manuales:** Formulario web de carga validada (si es necesario)

**Monitoreo:** - Dashboard de salud del sistema (última actualización, errores, registros procesados) - Alertas automáticas si falla proceso ETL

### 9.3 Evolución del Sistema

**Roadmap futuro (post-implementación):**

* **Corto plazo (6 meses):** Incorporar nuevos módulos según necesidades identificadas
* **Mediano plazo (1 año):** Mejorar integración con sistemas MINEDUC existentes
* **Largo plazo (2 años):** Expandir análisis y proyecciones según demanda institucional

## 10. PRÓXIMOS PASOS

### Para SEEMTP

1. **Semana 1:** Revisión y aprobación de propuesta por Coordinación
2. **Semana 2:** Gestión de presupuesto y aprobación financiera
3. **Semana 3:** Coordinación con TI MINEDUC para accesos y recursos
4. **Semana 4:** Llamado a contratación de desarrollador/a externo/a
5. **Semana 5:** Inicio Fase I del proyecto

### Para el Desarrollador Externo (una vez contratado)

1. **Día 1-3:** Reunión de kick-off, revisión de documentación, accesos
2. **Semana 1:** Coordinación con TI y levantamiento inicial
3. **Semana 2-4:** Integración de datos
4. **Semana 5-8:** Desarrollo de visualización y puesta en marcha

**Documento preparado:** Octubre 2024  
**Versión:** 2.0  
**Estado:** Propuesta para revisión interna