

# **Levantamiento de requerimientos para el diseño de una infraestructura computacional para la observación del cambio climático**

**Resumen Ejecutivo**

**Diego Rivera, Universidad del Desarrollo**

**Mario Lillo, Universidad de Concepción**

**Santiago, noviembre de 2021**

## Resumen Ejecutivo

Contexto: Este documento integra y sintetiza los principales aspectos, directrices y conclusiones abordadas en los informes de la asesoría sobre un sistema informático y su gobernanza, que sirva de soporte al Observatorio de Cambio Climático (OCC).

Las condiciones en las que se desarrollará este OCC son: disponibilidad de múltiples fuentes de datos públicas y privadas, heterogeneidad de formatos, soportes, frecuencias de actualización de variables, heterogeneidad en las capacidades y los desarrollos informáticos de cada una de las instituciones y la falta de una gobernanza que coordine los diferentes generadores y consumidores de datos.

Proponemos un producto mínimo viable basado en la información desarrollada en el AR6 en una plataforma liviana que considere tanto las variables de Chile continental y océano. En esta plataforma se podrán generar procesamientos a manera del *Climate Analysis and Plotting Tools*<sup>1</sup>. Seguidamente será posible incorporar nuevas variables y datos generados en terreno.

El OCC, como una herramienta de libre acceso con una gobernanza asociada a un ministerio y no a instituciones específicas, elimina barreras de entrada de los equipos científicos nacionales, de los alumnos de postgrado de diferentes instituciones y de la comunidad en general.

Este resumen se estructura sobre cuatro elementos principales:

- a. Estándares y protocolos
- b. Gobernanza
- c. Variables Fundamentales
- d. Barreras de entrada, tanto tecnológicas como de gobernanza de los datos.

---

<sup>1</sup> Climate Analysis and Plotting Tools

<https://psl.noaa.gov/cgi-bin/data/getpage.pl>

La misión del NOAA Physical Sciences Laboratory (PSL) es llevar a cabo investigaciones científicas para observar, comprender, modelar, predecir y pronosticar el tiempo, el agua y los extremos climáticos y sus impactos.

Es una plataforma de análisis y visualización de datos de clima histórico y futuro. Es un buen ejemplo de disponibilidad herramientas que bajan las barreras de entrada a investigadores jóvenes y estudiantes. Provee de un sitio en el cual se consolidan y actualizan diferentes fuentes de datos. Permite realizar comparaciones y operaciones complejas

FACTS - FACility for Weather and Climate Assessment

<https://psl.noaa.gov/repository/facts/>

Mantenido por el NOAA Physical Sciences Laboratory (PSL) permite comparar diferentes simulaciones de cambio climático futuro. Es un buen ejemplo de herramientas de comparación y acceso rápido a simulaciones de cambio climático.

## Tecnología y Estándares<sup>2</sup>

1. Desde una perspectiva tecnológica, la implementación de una estructura informática (cyberestructura) que albergue al OCC requiere de cinco etapas principales:

A. Motor de Integración: Sistema local (en cada institución) o centralizado, que permite gestionar los datos de las diferentes fuentes de cara a limpiar los datos (datos espurios) y detección de condiciones anormales. Esto es el primer paso para asegurar una interoperabilidad.

B. Factoría de Datos: Servicio centralizado que permite la Extracción/Transformación /Carga (ETL) de datos en un data lake.

C. Data Cube o Databrick: Corresponde a una matriz de datos de múltiples dimensiones. Permiten organizar los datos, simplificando su gestión mejorando el rendimiento de las consultas y el análisis. Es una buena solución para la gestión de datos espacio-temporales.

D. Analítica de datos: Esta etapa es la que agrega valor a los datos, aplicando inteligencia sobre ellos para contestar preguntas concretas.

E. Visualización y disposición de datos a los usuarios: Esta etapa considera reportes generales, específicos de información y acceso mediante APIs para la obtención de datos.

2. Estas etapas se pueden abordar tanto desde una aproximación de código abierto (*open source*), como desde soluciones empaquetadas. Como criterio de selección de una de las dos alternativas, se recomienda considerar variables como: Cyber-seguridad, escalabilidad del sistema, estandarización y compatibilidad con nuevos desarrollos, tiempo de desarrollo de nuevos productos.

De acuerdo con los criterios anteriores es recomendable el desarrollo de este tipo de infraestructura desde una aproximación de soluciones empaquetadas.

3. El atlas de cambio climático del AR6 se ha convertido en *gold standard* para el desarrollo de plataformas en tanto su concepción, despliegue de los datos y cumplimiento con los protocolos FAIR -*Findable, Accesible, Interoperable, Reusable*-.

Recomendamos utilizar esta plataforma como una guía para un producto final asociado al OCC.

4. No es necesario establecer nuevos estándares para el OCC. Lo anterior debido a que existen ya protocolos y estándares en uso que han mostrado éxito. Al respecto la Organización Meteorológica Mundial (WMO, por sus siglas en inglés) ha hecho posible el intercambio internacional de datos mediante el establecimiento de normas, reglamentos y prácticas para garantizar la sincronización e interoperabilidad a nivel global.

5. Se recomienda establecer protocolos de intercambio y usos de datos usando estándares ya probados. En este punto, la Dirección Meteorológica, dada la participación en la Organización

---

<sup>2</sup> Mayores detalles de la ciber-infraestructura, incluyendo una propuesta de actividades, en el Reporte 6 “Propuesta de una arquitectura computacional que contenga información asociada al monitoreo de variables relevantes para el cambio climático: atmosféricas, oceánicas, terrestres, sociales, entre otras” § 20 bis - § 33

Meteorológica Mundial es un actor clave para transferir experiencias al resto del aparato público.

6. Puesto que se requiere coordinar diferentes fuentes de datos y diferentes formatos, establecer protocolos e institucionalidad que permita la interoperabilidad; esto se puede definir como **un problema de gobernanza y no a un problema tecnológico**<sup>3</sup>. El enfoque de la WMO es un caso exitoso de **sistema distribuido con gobernanza fuerte**.

### **Gobernanza**

7. Se deben desarrollar las capacidades mínimas para que permitan asegurar la calidad de los datos, garantizar la protección y gestionar el ciclo de vida de la información.

Actualmente<sup>4</sup>, a partir de una encuesta a instituciones públicas<sup>5</sup> relacionadas con la generación y uso de datos, sólo 5 de 10 instituciones declaran la existencia e implementación de una política de datos y sólo 6 de 10 instituciones declaran adherir a algunos o todos los principios FAIR.

8. Actualmente, no existe una gobernanza de datos a nivel que implemente estándares, que establezca y haga cumplir prácticas adecuadas, roles y responsabilidades de cada institución o entidad que genera datos.

9. Es recomendable implementar una institucionalidad que coordine a los organismos generadores y consumidores de datos, defina las reglas (políticas, normas, directrices, protocolos), formas de decisión, responsabilidad y métodos de ejecución que rijan un sistema generador y gestor de datos.

El proceso de diseño e implementación de una gobernanza<sup>6</sup> no se puede desacoplar del diseño e implementación de la infraestructura. Por ejemplo, la propuesta de un sistema distribuido<sup>7</sup> de generadores y usuarios de datos, requiere de una gobernanza fuerte y centralizada.

---

<sup>3</sup> Mayores detalles en Ladley, J. (2019). Data governance: How to design, deploy, and sustain an effective data governance program. Academic Press.

<sup>4</sup> Mayores detalles en Reporte 5 “Reporte 6 Bases para el dimensionamiento del Observatorio de Cambio Climático que incluya aspectos de capacidades institucionales, recursos humanos e infraestructura y Recomendaciones para que las instituciones públicas permitan la interoperabilidad de sus sistemas de información” § 8

<sup>5</sup> Agroseguros, Instituto de Fomento Pesquero, Ministerio de Bienes Nacionales (SNIT), CIREN, Servicio Hidrográfico y Oceanográfico de la Armada (SHOA), Corporación Nacional Forestal (CONAF), INIA Chile, Dirección Meteorológica de Chile, Servicio Aerofotogramétrico de la FACH, Dirección de Obras Hidráulicas e Instituto Antártico Chileno.

<sup>6</sup> Mayores detalles en Reporte 5 “Reporte 6 Bases para el dimensionamiento del Observatorio de Cambio Climático que incluya aspectos de capacidades institucionales, recursos humanos e infraestructura y Recomendaciones para que las instituciones públicas permitan la interoperabilidad de sus sistemas de información” § 14 - § 20

<sup>7</sup> El principal beneficio de una arquitectura de este tipo es entregar a los diferentes perfiles de usuario acceso directo a los datos e información que sistemáticamente se dejará disponible para su uso, y a la agregación de valor que estos usuarios le entregarán mediante el uso y desarrollo de aplicaciones orientadas al manejo de datos asociados al agua, teniendo un impacto importante en organismos públicos y privados

10. En conclusión, el Observatorio de Cambio Climático -OCC- es una gobernanza que, a través de una infraestructura, habilitará la exposición y explotación de datos de interés científico. Así, el OCC deviene en una institucionalidad de coordinación que busca apoyar la generación de talento y soluciones para la economía digital basada en datos.

#### **Variables fundamentales<sup>8</sup>**

11. El Comité Científico de Cambio Climático indica la importancia de los ciclos del agua y carbono a nivel de cuencas hidrográficas. Al respecto, el *Global Climate Observing System* estableció los Indicadores Climáticos Globales como un conjunto de parámetros que describen el clima sin reducir el cambio climático únicamente a la temperatura<sup>9</sup>.

12. Las *Essential Ocean Variables* -EOV- son definidas por el *Global Ocean Observing System* GOOS que es liderado por el *Intergovernmental Oceanographic Commission* (IOC), parte de la UNESCO.

13. El conjunto mínimo de variables propuesto a incluir en el OCC es: precipitación, nivel de aguas subterráneas, caudal en desembocadura, altura superficial del mar, contenido de calor del océano, acidez del océano y carbono atmosférico.

#### **Barreras de entrada**

14. Las barreras de entrada tecnológica a abordar son:

- a. Falta de una normativa local, con alcance internacional, en cuanto a la generación y disposición de datos. Esto implica establecer una rectoría técnica que establezca requerimientos mínimos y coordine los procesos para informar y compartir datos. En un sistema distribuido, los generadores de datos retienen sus obligaciones de control de calidad y aseguramiento de calidad de los datos.
- b. Heterogeneidad en las capacidades y los desarrollos informáticos de cada una de las instituciones, puesto que los organismos generadores de datos no han implementado políticas de datos que consideren los principios FAIR.
- c. Heterogeneidad en las estructuras de datos y vocabulario utilizado (falta de un diccionario común)
- d. Falta de responsables en cada institución (*Data Chief Officer*) de los datos<sup>10</sup>.

15. Las barreras de entrada científicas a abordar son:

- a. Actualmente, tanto la capacidad de cálculo, como la capacidad de acceso y almacenamiento de datos relacionados con la observación del cambio climático se encuentran en instituciones específicas. Ello deviene en una barrera de entrada para otros grupos y organizaciones.
- b. Se requiere trabajar una gobernanza de datos que propenda a la democratización de los datos. El reciente convenio con Copernicus puede ser un piloto.

---

<sup>8</sup> Mayores detalles en Reporte 5 “Reporte 6 Bases para el dimensionamiento del Observatorio de Cambio Climático que incluya aspectos de capacidades institucionales, recursos humanos e infraestructura y Recomendaciones para que las instituciones públicas permitan la interoperabilidad de sus sistemas de información” § 9 - § 13, Tabla 1.

<sup>9</sup> Las variables son: Temperatura y energía (Temperatura superficial y Contenido de calor del océano), Composición atmosférica (Concentración de CO<sub>2</sub> atmosférico), Agua y océanos (Acidificación y nivel del mar) y Criósfera (Glaciares y Extensión del hielo en la Antártica y el Ártico).

<sup>10</sup> Ver Carruthers, C., & Jackson, P. (2020). The chief data officer's playbook. Facet Publishing y <https://www2.deloitte.com/us/en/insights/industry/public-sector/chief-data-officer-government-playbook.html>