

Adopción de plataforma de Big Data en el Banco Central de Chile

Viviana Rosales Salinas
División de Estadísticas

Agenda

- I. Antecedentes
- II. Introducción
- III. ¿Por qué es relevante el uso de los documentos tributarios electrónicos?
- IV. Objetivos y alcance del proyecto de implementación de una plataforma de Big Data
- V. Requerimientos
- VI. Implementación

A globe on a stand is centered in the image. The globe is semi-transparent, showing the continents. It sits on a dark, circular base. The background is a dark blue grid pattern. The word "Antecedentes" is written in white, sans-serif font, positioned to the right of the globe. A small, vertical gold bar is located at the end of the word.

Antecedentes |

Antecedentes

El 1º de enero de 2018 el Servicio de Impuestos Internos (SII) de Chile completó con éxito la implementación obligatoria de la Factura Electrónica, con la incorporación de las empresas rurales. Este proceso se había iniciado en 2014 con la adhesión obligatoria de las grandes compañías que operan en el país.

Esto significó que, a partir de esa fecha, todas las empresas del país deben reportar al SII sus transacciones con otras empresas de manera electrónica, detallando los bienes o servicios transados y sus respectivos impuestos, entre otros detalles.

Por otra parte, el 1º de abril de 2021 entró en vigencia la obligatoriedad de la emisión electrónica de las Boletas de venta de bienes y servicios.



Introducción

Introducción

El Banco Central de Chile, para llevar a cabo su labor de productor de las Cuentas Nacionales y para realizar aquellos análisis que contribuyen a mantener la estabilidad de la moneda, recibe y procesa registros administrativos provenientes de diversas fuentes.

A partir de octubre de 2019 el Banco comenzó a recibir los Documentos Tributarios Electrónicos (DTEs), en particular la Factura Electrónica, que son una fuente de información masiva y corresponden a la base de registros administrativos más grande que se haya recibido.

Este hecho generó la necesidad de contar con herramientas que permitan no solo almacenar grandes volúmenes de datos, sino que además hagan posible su explotación.





¿Por qué es relevante el uso de los documentos tributarios electrónicos?

¿Por qué es relevante el uso de los documentos tributarios electrónicos?

Los datos contenidos en la Factura Electrónica presentan una oportunidad única de acceso a información sobre las transacciones intermedias en bienes y servicios de la economía.

Mientras que los datos presentas en la Boleta Electrónica nos permiten conocer los detalles de las ventas a los consumidores finales.

BSALE DEMO INTEGRACION		RUT: 1-9 FACTURA ELECTRÓNICA			
 ARRIENDO DE SOFTWARE La Quebrada 1005 Puerto Varas Puerto Varas Mail: ebahamonde@imaginex.cl Teléfono: +56225847206		N° 5797 S.I.I. -			
SEÑOR:	Maquinarias Express	FECHA EMISIÓN:	01/01/2020		
RUT:	88.888.888-8	FECHA VENCIMIENTO:	01/01/2020		
GIRO:	Arriendo de maquinaria pesada	CIUDAD:	PUERTO MONTT		
DIRECCIÓN:	Los Alamos #122	FORMA DE PAGO:	TARJETA DEBITO		
COMUNA:	Puerto Montt				
Codigo	ITEM	VALOR UNITARIO	CANTIDAD	% dto.	SUBTOTAL
157978039593108	Toma requerimientos (3 rns)	\$ 37.815	1	0.0 %	\$ 37.815
157978041393109	Implementación etapa inicial	\$ 168.067	1	0.0 %	\$ 168.067
157978043393110	Horas desarrollo	\$ 25.008	16	0.0 %	\$ 400.128
157978049393111	Pruebas SET BASICO	\$ 8.395	1	0.0 %	\$ 8.395
157978053493112	Implementación final (sin mms)	\$ 100.840	1	0.0 %	\$ 100.840

 DON MATÍAS CARNES		RUT 76.839.883-4 BOLETA ELECTRONICA N° 8 S.I.I. - PUERTO VARAS		
CARNES DON MATIAS				
Giro:	MAYORISTAS DE CARNES			
Dirección:	LA QUEBRADA 1005			
Comuna:	PUERTO VARAS			
Ciudad:	PUERTO VARAS			
Fono:	229381115			
Vendedor:	Juan Tapia			
Fecha:	22/07/2020			
Forma de Pago:	EFECTIVO			
Sucursal:	LA QUEBRADA 1005, PUERTO VARAS			
CANT.	ÍTEM	VALOR U.	DESC.	SUBTOTAL
5	LOMO VETADO KG	\$ 8.999	\$ 0	\$ 44.995
NETO:		\$ 37.811		
IVA:		\$ 7.184		
Total Exento:		\$ 0		
Vuelto:		\$ 0		
TOTAL:		\$ 44.995		

¿Por qué es relevante el uso de los documentos tributarios electrónicos?

Los desgloses y el volumen de la información disponible abren la puerta a la compilación de nuevas y mejores estadísticas, así como también de nuevas líneas de investigación.

Su uso permitirá la creación de nuevos indicadores coyunturales para la toma de decisiones de política monetaria (inversión, precios), además de la generación de indicadores líderes para determinar cambios en el ciclo económico a partir de la información de transacciones entre empresas.

La información se encontrará disponible con un desfase mensual, lo que permite tener una visión oportuna de las relaciones económicas antes mencionadas.

¿Por qué es relevante el uso de los documentos tributarios electrónicos?

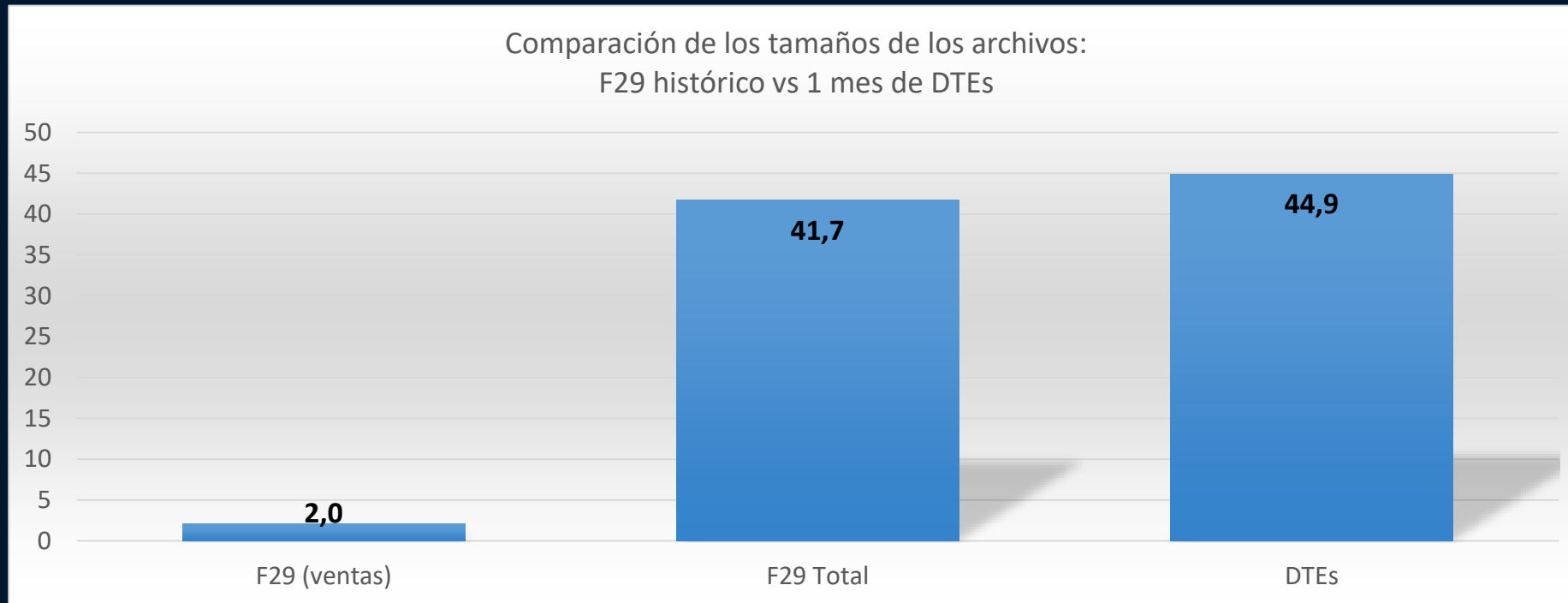
Si bien se ha estado procesando esta gran cantidad de información en nuevos servidores, tanto la capacidad de almacenamiento como de procesamiento sigue siendo un problema:

- Los análisis requeridos en la elaboración de las Cuentas Nacionales necesitan tener una perspectiva de series de tiempo, por lo que es necesario “levantar” toda la historia de datos en simultáneo, lo que no es posible con las tecnologías actuales.
- No ha sido posible la ejecución de modelos complejos utilizando toda la data.
- Ha sido necesario comprar más capacidad de almacenamiento de manera recurrente.


¿Por qué es relevante el uso de los documentos tributarios electrónicos?

1 mes de Factura Electrónica implica:

1. 40 GB de memoria en formato .txt
2. Más de 200 millones de registros únicos en la base de datos
3. 31 millones de relaciones comerciales



Fuente: elaboración propia



Objetivos y alcance del proyecto de implementación de una plataforma de Big Data

Objetivos

Mejorar las estadísticas actuales, crear nuevos indicadores y fortalecer la investigación, ayudando a la toma de decisiones de política monetaria.

Aprovechar la sinergia entre las unidades para realizar una explotación centralizada y eficiente de la información, evitando duplicaciones.

Alcances

Habilitación de una plataforma (infraestructura + software) que permita el almacenamiento, depuración y análisis de grandes volúmenes de información.

Implementación de dos casos de uso, uno asociado a la Factura Electrónica y otro a la Boleta Electrónica, que contemplen la carga de los datos, el desarrollo de modelos, la depuración de la información y la elaboración de reportería.

Adquisición de competencias para operar con la plataforma, tanto desde el punto de vista de análisis de negocio como de desarrollo TI y administración.

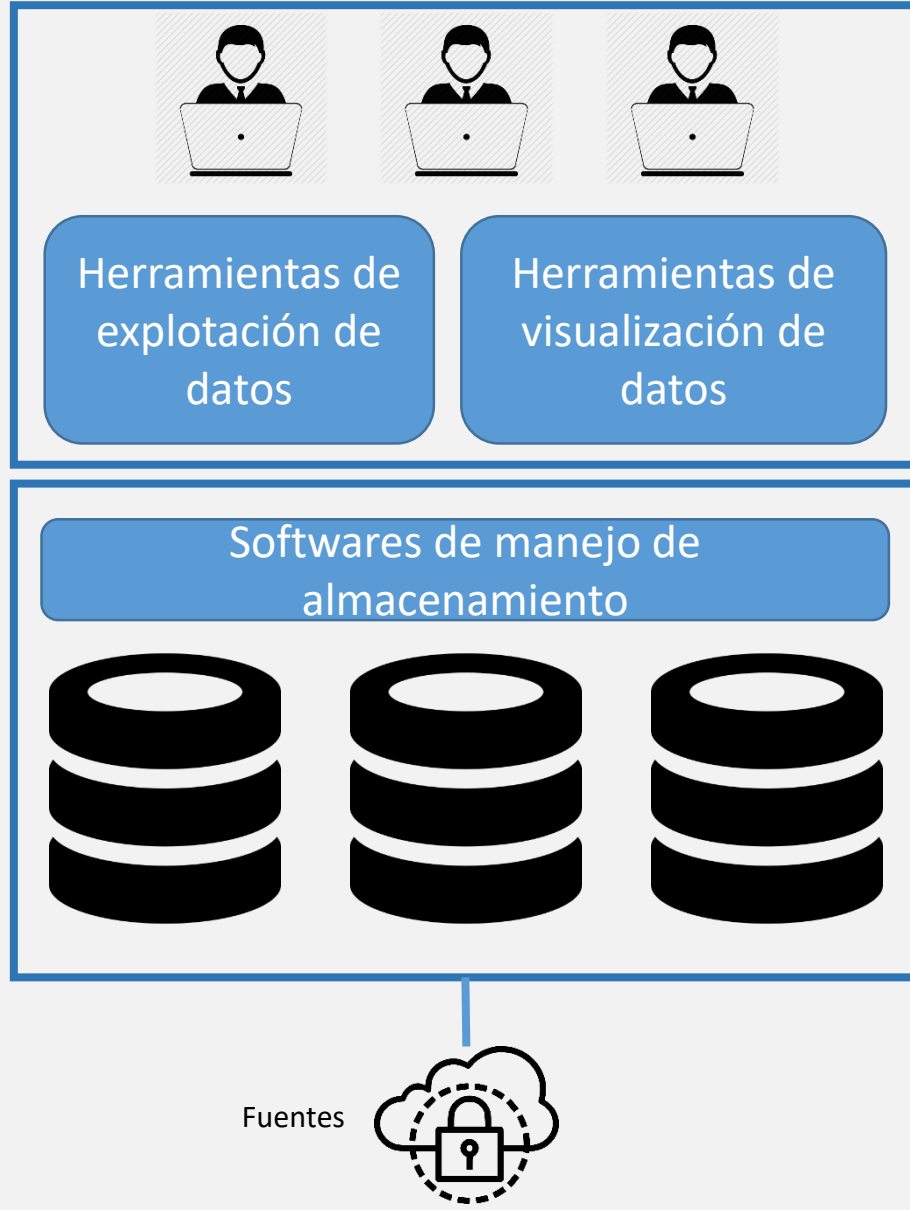


Requerimientos |

16 18 20 22 24 26

Requerimientos

- Plataforma on-premise (debido a restricciones de la fuente de datos).
- Incluir herramientas para que los usuarios puedan construir y configurar integraciones entre distintos ambientes de manera autónoma.
- Ser capaz de realizar la carga de datos de manera distribuida en tiempos acotados.
- Permitir la ejecución de scripts de R y Python utilizando procesamiento distribuido, en particular, algoritmos de clasificación de texto.
- Permitir la creación de Data Labs.
- Permitir la creación de reportería.
- Permitir la conexión con herramientas actualmente en uso en el Banco.
- Permitir el perfilamiento a distintos ambientes de datos (tablas, scripts, procedimientos almacenados, etc).
- Capacitación para todo tipo de usuarios en el uso de la Plataforma.



Casos de Uso:

- Elaboración de procesos para la extracción, carga y procesamiento de:
 - Factura electrónica y Boleta electrónica
 - Referencias (que permite vincular notas de crédito con FE y BE)
- Ejecución de consultas en sentencias SQL con el total de la información histórica para realizar agrupaciones, uniones entre múltiples tablas, búsqueda por variables relevantes, etc.
- Ejecución de una serie de procesos en R y Python (proporcionados por el Banco) con un enfoque de *parallel computing*.
- Elaboración de Data Mart con los microdatos y el resultado de la clasificación de texto.
- Elaboración de reportería.

A glass globe on a stand, centered on the map of South America. The globe is set against a dark blue background with a faint grid pattern. The word "Implementación" is written in white text across the globe, followed by a vertical gold bar.

Implementación |

Implementación

- Julio 2021: adjudicación de la licitación
- Primer trimestre 2022: implementación del Caso de Uso de la Factura Electrónica
- Tercer trimestre 2022: implementación del Caso de Uso de la Boleta Electrónica

Adopción de plataforma de Big Data en el Banco Central de Chile

Viviana Rosales Salinas
División de Estadísticas