

Manual para la actualización de la BDD del DIEE utilizando las matrices de reglas y prioridades

Octubre, 2018

Contenido

Antecedentes	3
Objetivo	3
1. Matriz de Prioridades	3
2. Matriz de reglas.....	7
Anexos	11

Manual para la actualización de BDD del DIEE, utilizando las matrices de Reglas y Prioridades

Antecedentes

El directorio de empresas es una operación estadística que se construye a partir de información de registros administrativos procedentes de fuentes como el SRI, IESS, fuentes internas de las encuestas que levanta el INEC y del sistema de Call Center implementado por el DIEE. Con la finalidad de mejorar la calidad de la información de las variables del Directorio se han considerado otras fuentes externas como: el Ministerio de Educación, Superintendencia de la Economía Popular y Solidaria (SEPS), Servicio Nacional de Contratación Pública (SERCOP), entre otras.

El contar con diversas fuentes que aportan con información de variables similares para la creación de la base del Directorio de Empresas y Establecimientos, es necesario determinar que variables se deben seleccionar de estas fuentes; por lo tanto en la unidad del DIEE se diseñaron matrices en las que se describen, que procedimientos se debe tomar en cuenta antes, durante y después del procesamiento para la actualización de la BDD del DIEE.

El presente manual detalla los procedimientos a seguir al momento de actualizar la Base de Datos del Directorio de Empresas y Establecimientos, utilizando las matrices de reglas y prioridades.

Objetivo

Contar con un manual para la utilización de la matriz de reglas y prioridades, el actualizar la BDD del DIEE.

1. Matriz de Prioridades

La matriz de prioridades se diseñó en base a la información que aportan las diferentes fuentes al Directorio de empresas y tiene las siguientes características:

No. PRIORIDAD	UNIDAD	VARIABLE	FUENTE	PRIORIDAD
---------------	--------	----------	--------	-----------

No. Prioridad: Conteo de las prioridades diseñadas según su necesidad.

Unidad: Describe a qué nivel se debe aplicar la prioridad, la cual puede ser: Empresa, unidad legal o unidad local.

Variable: En este campo se define a que variable se debe aplicar la prioridad, y pueden ser:

N°	Variables	N°	Variables
1	actividad_economica	18	id_actividad_comercio_exterior
2	actividad_secundaria	19	id_acto_juridico
3	coordenadas	20	id_clase_contribuyente
4	dirección	21	id_ruc

5	empleo	22	id_zona_sector_manzana
6	fecha_acto_juridico	23	nombre_comercial
7	fecha_actualizacion	24	numero_unidad_local
8	fecha_apertura	25	obligado_llevar_contabilidad
9	fecha_cese_actividad	26	razon_social
10	fecha_cierre	27	ruc_acto_juridico
11	fecha_inicio_actividad	28	ruc_adscrita
12	fecha_inscripcion	29	unidad_legal_estado
13	fecha_reinicio_actividad	30	unidad_legal_tipo
14	fecha_suspension	31	unidad_local_estado
15	Forma institucional	32	unidad_local_nombre
16	forma_institucional	33	ventas
17	geografia		

Fuente: En este casillero se abrevia las iniciales de la institución que provee la información

No.	DESCRIPCIÓN	ABREVIATURA	No.	DESCRIPCIÓN	ABREVIATURA
1	INDUSTRIALES	IND	13	ANÁLISIS DIEE	AND
2	AMBIENTALES	AMB	14	SUPER DE COMPAÑÍAS	SCI
3	ACTI	ACT	15	EXHAUSTIVA	EXH
4	CALL CENTER	CAL	16	GEORREFERENCIA	GEO
5	SRI	SRI	17	NOMENCLATURAS	NOM
6	IESS	IES	18	LABORATORIO DE DINÁMICA LABORAL	LDL
7	ENEMDU	ENM	19	SUPER DE BANCOS	SBS
8	CENEC	CEN	20	INA-R	INR
9	SISMO	SIS	21	IMPUTACIONES	IMP
10	CARTOGRAFÍA - CAMPO	CAM	22	SERCOP	SER
11	CARTOGRAFÍA - CALL CENTER	CAC	23	MINISTERIO DE EDUCACIÓN	MED
12	SIPRO	SIP	24	SUPERINTENDENCIA DE ECONOMÍA POPULAR Y SOLIDARIA	SEP

Prioridad: En esta columna se da la prioridad con la que se utilizará la información y los rangos son de 1 a n (n orden de prioridad), donde 1 es la prioridad.

Ejemplo: Si se debe actualizar código de CIU 4.0 en una unidad legal, y se dispone información del código CIU 4.0 del Ministerio de Educación (MED) e información de Cartografía-Campo (CAM), la que predomina es la información del MED

No. PRIORIDAD	UNIDAD	VARIABLE	FUENTE	PRIORIDAD
6	Unidad Local	actividad_economica	MED	1
8	Unidad Local	actividad_economica	CAM	2

Con la gama de fuentes de información se puede realizar diferentes combinaciones

según la prioridad dada a cada una de ellas y con ello se actualizó la información del Directorio de Empresas y Establecimientos.

1.1. Como utilizar la matriz de prioridades

La matriz será utilizada por el analista de procesamiento, quien es el encargado de actualizar la BDD del DICE que luego será objeto de difusión.

La utilización de esta herramienta (matriz de prioridades), le permite discriminar desde el proceso de procesamiento las fuentes que actualizaran las diferencias variables con las que esta construida la BDD del DICE y pasará a nombrarse como nueva para el año de publicación.

La Actualización de la BDD del DICE, deberá hacerse por variables con la utilización de las diferentes fuentes disponibles, Tal como se muestra en la siguiente tabla:

No. PRIORIDAD	UNIDAD	VARIABLE	FUENTE	PRIORIDAD
110	Empresa	actividad_economica	ACT	1
111	Empresa	actividad_economica	AMB	1
112	Empresa	actividad_economica	CEN	1
113	Empresa	actividad_economica	EXH	1
114	Empresa	actividad_economica	IND	1
115	Empresa	actividad_economica	MED	1
116	Empresa	actividad_economica	MTU	2
117	Empresa	actividad_economica	CAM	2
118	Empresa	actividad_economica	AND	3
119	Empresa	actividad_economica	CAL	3
120	Empresa	actividad_economica	LDL	3
121	Empresa	actividad_economica	NOM	3
122	Empresa	actividad_economica	MC2	4
123	Empresa	actividad_economica	MC4	4
124	Empresa	actividad_economica	MCO	4
125	Empresa	actividad_economica	MEC	4
126	Empresa	actividad_economica	SRI	4
127	Empresa	actividad_economica	SIP	4
128	Empresa	actividad_economica	SCI	5

Como se puede observar, para actualizar la actividad económica las fuentes que tienen prioridad son: Encuesta nacional de Actividades de Ciencia Tecnología e Innovación (ACT); Estadísticas Ambientales ABM... y en segundo grado de prioridad están las: Ministerio de Turismo MTU, cartografía-Campo CAM....; ese es el orden con el cual se irá actualizando la información para la variable actividad económica; adicional a ello, además de seleccionar la fuente también deberá seleccionar aquella fuente que contenga la fecha más actual.

1.2. Proceso para la actualización de fuentes de la matriz de prioridades.

Al existir diversas fuentes de información que contribuyen a la creación de directorio de empresas, es importante que se defina un proceso sobre el tratamiento que se

deba dar a las mismas al momento de asignar la prioridad que permita seleccionar la información necesaria para la construcción de la base de datos. Este proceso deberá considerar los siguientes puntos:

1.2.1. Identificación de las fuentes en el periodo de referencia

Identificar las fuentes comprende realizar un análisis comparativo entre la variable de actualización (Contactos, actividad económica, ubicación geográfica, etc) dadas las diferentes fuentes de información (IND, SIP, CALL, etc), para realizar este proceso el directorio de empresas deberá mantener los acuerdos para el intercambio de información con las productoras que puedan aportar con información oportuna, lo que permitirá la vinculación para el manejo de variables y establecer tiempos en base a cronogramas que permitan realizar los procesos de selección de información por parte de los analistas de DICE. Este proceso es primordial por cuanto de ello dependerá saber con qué fuentes se cuenta para iniciar con la selección de las mismas.

1.2.2. Actualización de fuentes en la matriz de prioridades

Consecuente al punto anterior es necesario que la matriz de prioridades se deba actualizar previo a la construcción del directorio de empresas y establecimientos, esto debido a que no todas las fuentes aportan con información continua, tal es el caso de fuentes como: Censo Nacional Económico-2010 CENEC (CEN), Encuesta Exhaustiva (EXH), entre otras. Si existieran nuevas fuentes que aporten información como el caso del Servicio Nacional de Contratación Pública SERCOP (SER) o el Ministerio de Educación MINEDUC (MED) que aportaron con información para la construcción del directorio 2016, y una vez que se haya analizado eficiencia y consistencia de la información, la fuente pasará a formar parte de la matriz de prioridades y por ser nueva fuente que aportará con información en periodo de referencia tendrá prioridad uno.

1.2.3. Omisión o eliminación de fuentes de información

Es importante mencionar que ninguna fuente de información es eliminada de la matriz de prioridades del Directorio de Empresas, esto debido a la priorización entre fuentes y las series históricas de las variables que se publican.

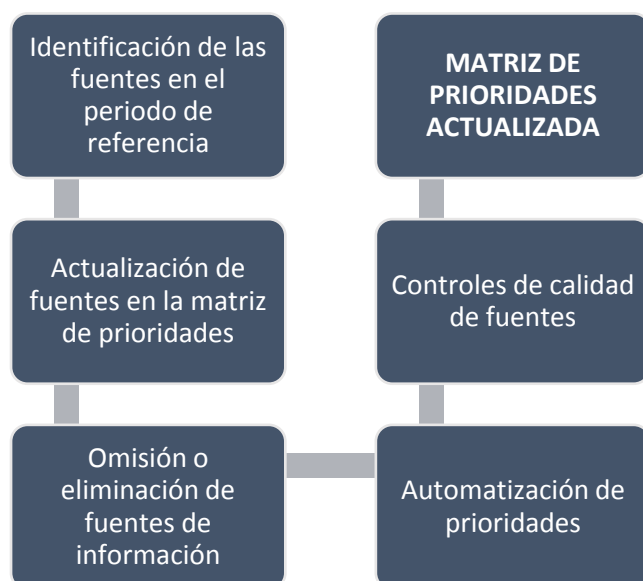
1.2.4. Automatización de prioridades

Identificado las fuentes de información que formaran parte de la matriz de prioridades, estas serán remitidas a procesamiento con la finalidad de que estas se incluyan en el proceso de automatización.

1.2.5. Controles de calidad de fuentes

Luego que en procesamiento realizaron la actualización de las diferentes variables, se procede a la revisión de las mismas en función de los criterios plasmados en la matriz de prioridades, con ello se identifica si algún registro fue o no actualizado.

Estos procesos se resumen en el siguiente diagrama:



2. Matriz de reglas

La matriz de reglas es un documento de guía para procesamiento, en el que se describe las acciones a seguir para el tratamiento de actualización de cada una de las variables que conforman la Base de Datos del Directorio de Empresas y Establecimientos.

La matriz de reglas esta conformada por los siguientes campos:

- **No. de Regla:** Conteo de reglas creadas.
- **Variables:** se detalla la variable a la cual se va aplicar la regla.

No. Variable	VARIABLE
1	Actividad de comercio exterior
2	Actividad Económica
3	Acto Jurídico
4	Clase contribuyente
5	Contacto
6	Continuidad
7	Dirección
8	Empleo Registrado
9	Empleo registrado y Plazas de Empleo Registrado
10	Empleo registrado-equivalente
11	Empleo registrado-Promedio
12	Empleo registrado-trimestral
13	Empleo registrado-Último
14	Estado
15	Estrato empleo
16	Estrato ventas
17	Estructura de Estratificación
18	Fecha actualización
19	Fecha cese actividad
20	Fecha inicio actividad
21	Forma institucional
22	Nombre comercial

No. Variable	VARIABLE
23	Número de unidades locales
24	Obligado a llevar contabilidad
25	Obligado llevar contabilidad
26	Origen de Estratificación
27	Personal afiliado
28	Plazas de Empleo Registrado
29	Plazas de Empleo Registrado - Último
30	Plazas de Empleo Registrado_Promedio
31	Plazas de Empleo Registrado-equivalente
32	Plazas de Empleo Registrado-trimestral
33	Remuneraciones
34	Situación
35	Tamaño de empresa
36	Teléfonos
37	Tipo de unidad legal
38	Tipo de unidad local
39	Ubicación Geográfica
40	Unidad legal estado
41	Universo
42	Variables de control para la geografía
43	Ventas

- **Fuente:** detalle de la fuente de donde proviene la información.

No. de Fuente	Fuente
1	Análisis DIEE
2	Empleo
3	ENEMDU
4	IESS
5	Industriales
6	Matriz de Prioridades
7	Ministerio de Educación
8	Ministerio de Turismo
9	Plazas
10	SERCOP
11	SRI, CALL CENTER
12	sueldo
12	Super de compañías
13	Superintendencia de Bancos

- **Origen:** Detalla el origen de creación de las reglas, producto de ingreso de nuevas fuentes o cambios en procesos metodológicos del DIEE.
- **Regla:** se detalla la solución ante el origen de una necesidad en las variables que conforman el DIEE.
- **Tipo de unidad:** En esta columna se define a qué nivel se debe aplicar la regla y puede ser: Empresas, empresas y establecimientos, establecimientos y unidad legal.
- **Documento de respaldo:** Se detallarán los documentos con los que se dio apertura a la regla, estos pueden ser mediante: email, acta de reunión entre otros, ejemplo: Informe de análisis de las juntas de agua.
- **Fecha de creación:** Se detalla la fecha en la que se creó la regla.
- **Fecha de actualización:** esta columna se utiliza en el caso de que la regla ya existe y solo se realiza una actualización de la misma.
- **Aplica:** En esta columna se ingresa SI o NO, por lo general es SI, pero hay ocasiones en las que se crea reglas nuevas para reemplazar completamente a una existentes, en estos casos se indica que cierta regla ya no se utilizar más; se selecciona el NO, con el fin de eliminar ninguna regla, con ello se preserva la historia de la BDD del DIEE.
- **Ejecución:** Cada regla tiene su tiempo de aplicación, por lo tanto este campo se describe cuando es su utilización; los mismos que pueden ser: al inicio, durante o al final del procesamiento, entre otros.

No.	Ejecución
1	Al final de cada procesamiento
2	Al final del procesamiento
3	Aplica para Análisis
4	Cuando se tenga información CENEC
5	Cuando se tenga información EXH
6	Durante el procesamiento
7	Durante el procesamiento (Ejecutado por Única vez)
8	En limpieza de información
9	Entrega de tabla consolidada de empresas
10	Entrega de tabla consolidada de establecimientos
11	Ocasional
12	Por única vez
13	Reprocesar hasta el 2009

- **Observación:** alguna descripción que no este contemplada en los campos antes descritos y que merezca ser tomada en cuenta.

Este documento es aplicado por los analistas y el área de procesamiento, ya que en conjunto y en función a las necesidades se las crea, sea por análisis a por procesamiento, con el objeto de obtener un proceso automatizado y así obtener cada vez un producto de mejor calidad.

La matriz de reglas esta diseñada a nivel de variable de esta manera tanto el área de análisis y procesamiento podrá identificar con facilidad la variable, la regla y el tipo de unidad el que se deberá aplicar. Por ejemplo: Se selecciona la variable ““Actividad de comercio exterior” en el formato se filtrará una regla según tipo de unidad de “Empresa”. El analista deberá considerar la regla para los registros en los que se aplique el origen.

2.1. Proceso para la actualización de la matriz de Reglas

Las reglas surgen de las necesidades del directorio de empresas y establecimientos, por tal motivo algunas se ajustan solamente a determinados casos he incluso en determinado tiempo, bajo este contexto es necesario definir un proceso que se debería adoptar al momento de crear las reglas que formarán parte de la matriz. Estos puntos se detallan de la siguiente manera:

2.1.1.Actualización de reglas

La actualización de reglas se realiza previa a la construcción del directorio de empresas y establecimientos, con la recopilación de novedades que se hayan presentado en la construcción del directorio anterior, debido a que dadas las experiencias se definirán las nuevas reglas y se mejorarán las vigentes, permitiendo identificar aquellas reglas que se aplicaran de manera continua.

2.1.2.Creación de reglas

Para la creación de la nueva regla se debe consideran lo siguiente: impacto de la regla dada las variables que intervienen, periodicidad, origen, unidad de aplicabilidad y el tipo de ejecución, que también se detalla en la matriz de reglas. Previo a creación

de toda regla es importante verificar que la regla o solo es necesario actualizar la existente.

2.1.3.Omisión de reglas

Debido a la naturaleza de las reglas, es importante que todas las reglas formen parte del historial de reglas, puesto que cada regla se creó con un fin, el cual puede ser replicado en procesos posteriores al del periodo de referencia.

2.1.4.Automatización de reglas

Toda regla o su mayoría, sea esta nueva o actualizada deberá cumplir con un proceso de automatización mediante Scripts u ETL's.

2.1.5.Control de calidad de reglas

Todas regla que haya sido actualizada o creada como nueva, deberá ser sometida a un proceso de control de calidad, bajo los criterios adoptados en el plan de control de calidad para todas las fases, para los cual siempre se deberá evaluar el indicador de correspondencia y de funcionalidad.

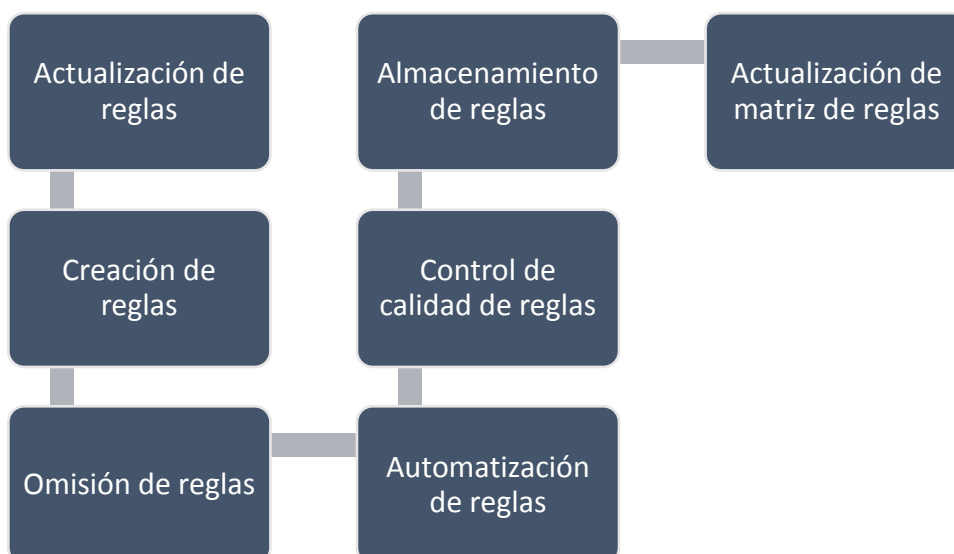
2.1.6.Almacenamiento de reglas

Todas las reglas se deberán almacenar en el archivo "Procesos_reglas_matriz" el cual se encuentra en formato txt, en este formato se evidenciará el código de la regla, el respaldo de la regla que identifica el motivo de la creación de la regla, y la automatización en lenguaje de programación (procesamiento).

2.1.7.Actualización de matriz de reglas:

Realizados los puntos anteriores se podrá contar con una matriz de reglas actualizada y funcional aplicable al directorio de empresas en procesamiento.

Los puntos detallados anteriormente se resumen en el siguiente diagrama:



2.2. Actualización de registros en la base de datos

Para cada actualización de información a la Base de Datos del DIEE, se debe utilizar la matriz de reglas, según variable de actualización.

2.2.1. Proceso

- Luego de la actualización de la información de la Base de Datos del DIEE, desde procesamiento, remitirá mediante correo electrónico los registros actualizados al área de análisis.
- Desde el área de análisis, informará que la Base de Datos del DIEE, para el año de publicación fue completamente actualizada y si fuere el caso también describirá las novedades encontradas.

Anexos

- 1. Matriz_reglas**
- 2. Matriz_prioridades**

Nombre y Apellido		Cargo
Elaborado por:	Oscar Romero	Analista
Revisado y Aprobado por:	Libertad Trujillo	Jefe de Gestión de Estadísticas Económicas en Base a Registros Administrativos

Información del Documento

Resumen: El presente manual contiene los antecedentes, objetivos y ordenamientos a seguir con respecto a la correcta aplicación de normas y procedimientos para la actualización de la BDD del DICE, con las matrices de reglas y prioridades.

CONTROL E HISTORIAL DE CAMBIOS

Versión	Descripción del cambio/ Autor (a)	Fecha de creación y actualización
1.0	Versión Original, Oscar Romero	14/02/2017
1.1	Actualización, según los nuevos formatos de matrices	24/10/2019