



Codificación del Censo de Población y Vivienda 2020

Etapas de la codificación



Actividades de preparación

Diseño de estrategias

Revisión de antecedentes

Elaboración de materiales

Desarrollo del sistema

Pruebas

Actualización de clasificaciones

Revisión de antecedentes

- Experiencia previa (algoritmos, flujos, sistemas, resultados)
- Revisión del estado del arte

Diseño de estrategias

- Por variable
- Diseño del Pre procesamiento
- Diseño de algoritmos de codificación

Elaboración de materiales

- Por variable
- Catálogos automáticos, tablas de acceso, tablas especiales, diccionarios de sinónimos
- Reglas de congruencia, rangos válidos y unicidad

Actualización de clasificaciones

- Por clasificación
- Reuniones con expertos
- Adecuación de descripciones, incluye y excluye
- Creación de nuevos grupos

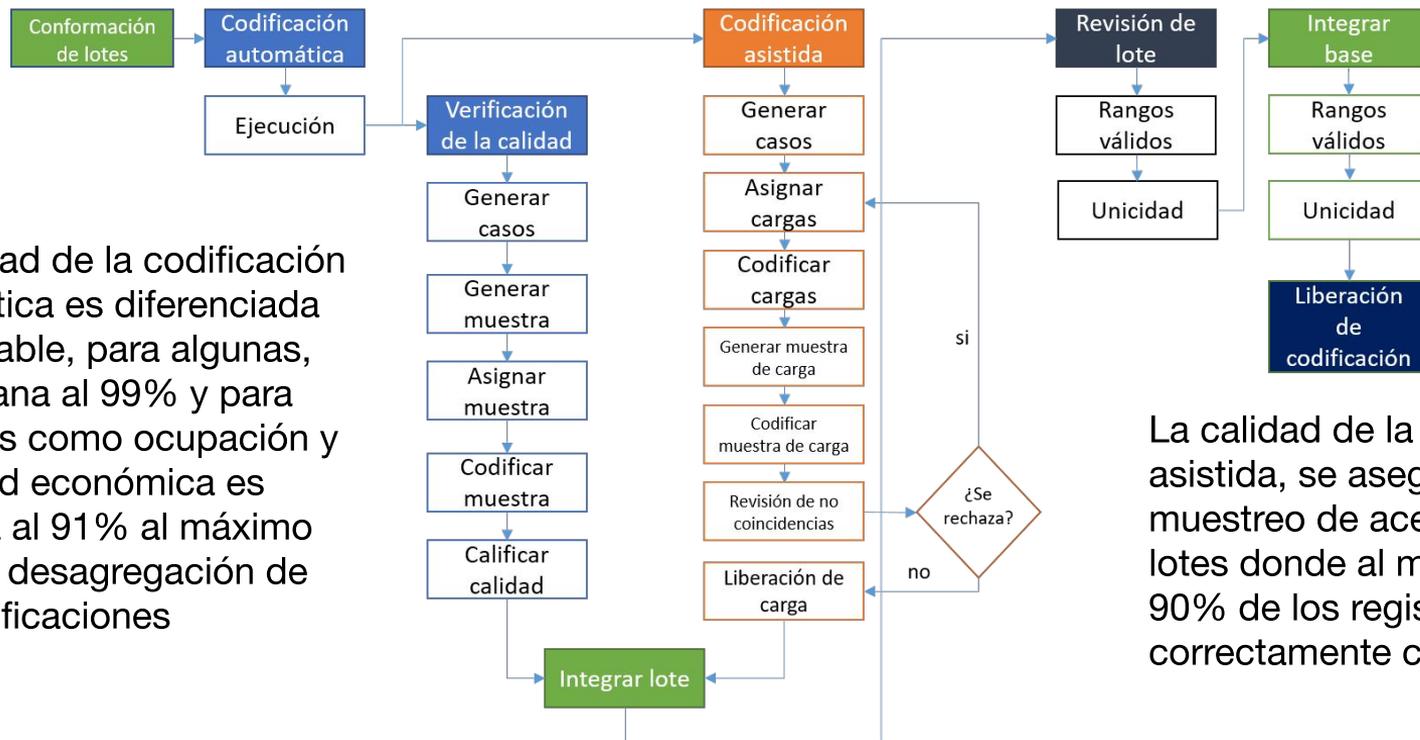
Desarrollo del sistema

- Requerimiento del sistema
- Programación de algoritmos
- Desarrollo de la interfaz

Pruebas

- Pruebas de codificación con información de censos anteriores
- Puesta a punto y liberación

Proceso de codificación



La calidad de la codificación automática es diferenciada por variable, para algunas, es cercana al 99% y para variables como ocupación y actividad económica es cercana al 91% al máximo nivel de desagregación de las clasificaciones

La calidad de la codificación asistida, se asegura con un muestreo de aceptación por lotes donde al menos el 90% de los registros son correctamente codificados

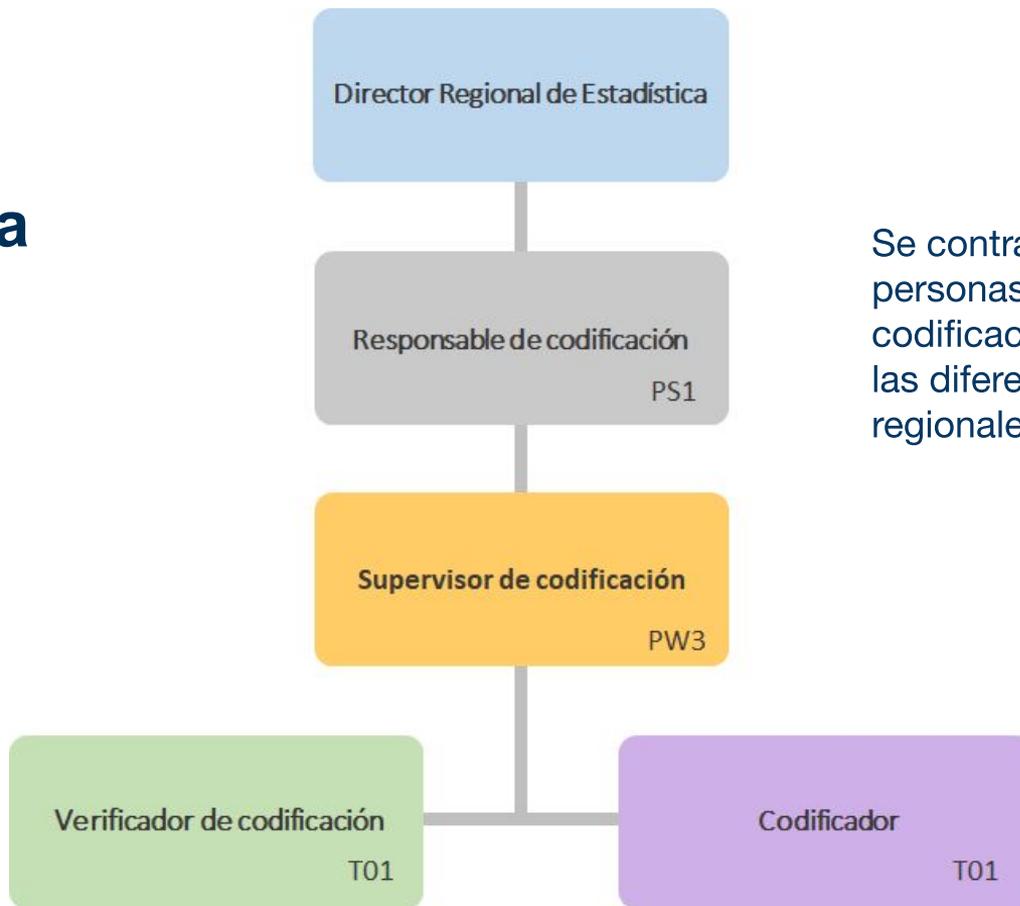
Codificación automática

- Corrección de palabras
- Lematización
- Truncado de palabras
(Stemming)

**Pre
procesamiento**

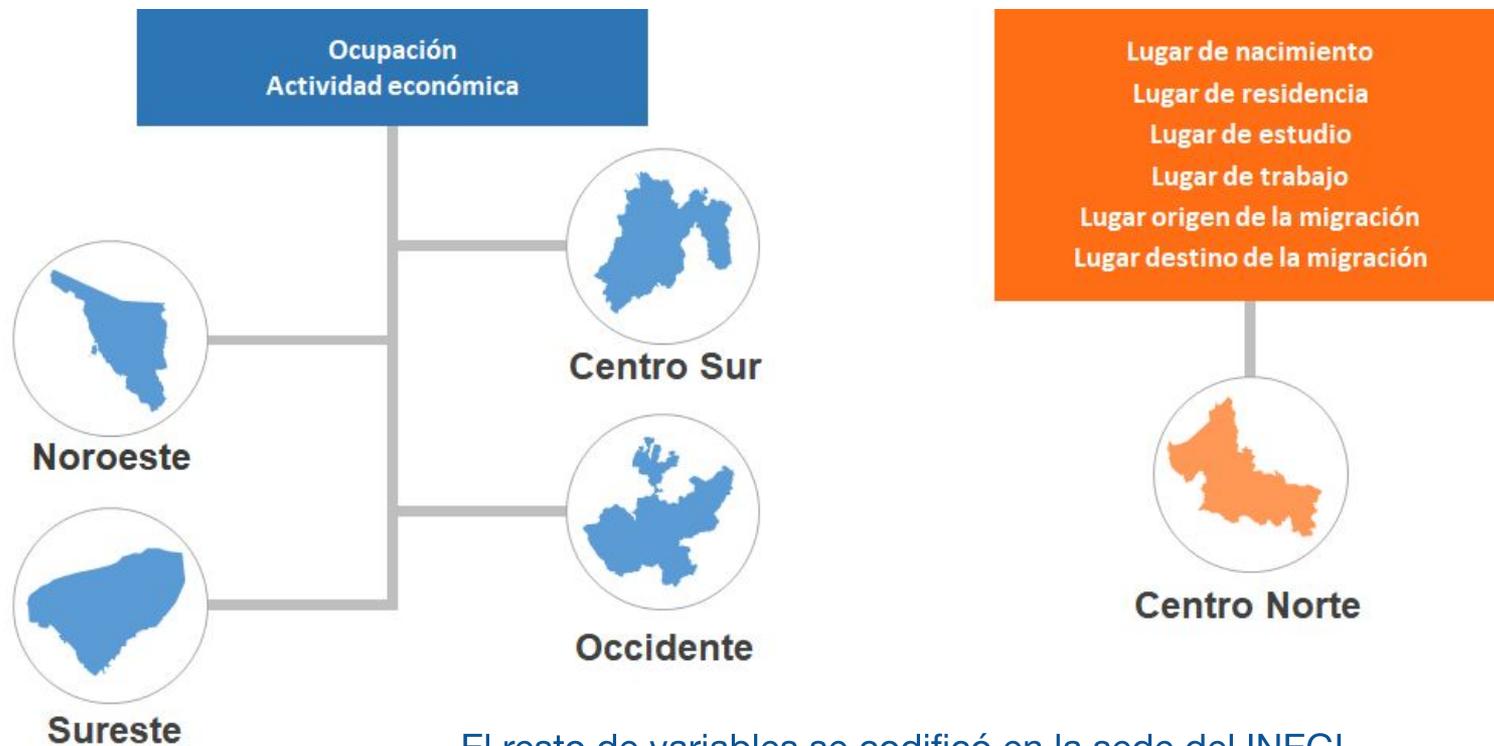
Ejecución de algoritmos determinísticos "If the text contains or is equal to ... then ..." diseñados de forma particular para cada variable

Organigrama regional



Se contrataron 357 personas para atender la codificación asistida en las diferentes sedes regionales

Regionales participantes



El resto de variables se codificó en la sede del INEGI

Sistema de codificación

Menú supervisor

Opciones de menú

- Codificación CPV2020
 - Codificación
 - Asignar carga
 - Supervisión
 - Generar muestra
 - Asignar muestra
 - Evaluar carga
 - Supervisar diferencias
 - Liberar carga
 - Seguimiento y monitoreo
 - Avance de cargas
 - Productividades
 - Seguimiento a cargas
 - Consulta
 - Entidades / Municipios
 - Países
 - SCIAN
 - SINCO
 - CMPE
 - Religión

Se tiene un sistema que permite administrar el proceso de codificación, tanto para la parte automática como asistida

Menú codificador

Bienvenido: CODIFICADOR APELLIDO_PAT APEL

Opciones de menú

- Codificación CPV2020
 - Codificación
 - Codificar
 - Consulta
 - Entidades / Municipios
 - Países
 - SCIAN
 - SINCO
 - CMPE
 - Religión

Codificación asistida

Variable: ACTIVIDAD ECONÓMICA Total de casos: 5 Casos faltantes: 5

Buscar caso Ir al caso faltante

Clave Actividad: Descripción:

Asignación de claves				Preguntas de apoyo			
Frecuencia	Caso	Ind Actividad	Actividad economica	Ocupacion	Posición en el trabajo	Servicios de Salud	Lugar de Trabajo
1	1	YY	A CHAPOLEAR	JORNALERO			
			Lugar de trabajo otro	Tareas	2-Subordinado		1-Campo agrícola
				CHAPOLEO			

BÚSQUEDA EN LA CLASIFICACIÓN ASISTIDA

Limpiar consulta

Cualquier palabra Frase exacta Todas las palabras

Duda conceptual Dudas anteriores

Clave Actividad Descripción

5413

Escribe aquí una descripción

Codificar

- 54-SERVICIOS PROFESIONALES, CIENTÍFICOS Y TÉCNICOS
 - 5413-SERVICIOS VETERINARIOS
 - 5413-1-BANCO DE ESPERMA ANIMAL
 - 5413-1-LABORATORIOS ANÁLISIS CLÍNICOS VETERINARIOS
 - 5413-SERVICIOS MÉDICOS VETERINARIOS PÚBLICOS O PRIVADOS

Resultados

Codificación por variable según registros codificados.

Fecha de elaboración: 22 de octubre de 2020.

Variable	Registros codificados	Codificación			Calidad Global
		Automática	%	Asistida	
Total	177,849,140	171,875,756	96.6	5,973,384	98.06%
Ocupación	5,601,641	3,915,829	69.9	1,685,812	93.31%
Actividad económica	5,601,641	4,278,062	76.4	1,323,579	91.98%
Religión	119,642,121	118,861,800	99.3	780,321	99.90%
Parentesco	482,700	472,259	97.8	10,441	99.99%
Lengua indígena	7,318,891	7,202,511	98.4	116,380	99.96%
Carreras	1,639,420	1,551,567	94.6	87,853	97.28%
Lugar de trabajo (País)	2,572	2,156	83.8	416	99.98%
Lugar de trabajo (Estado)	1,015,114	942,534	92.9	72,580	97.70%
Lugar de trabajo (Municipio)	1,015,114	942,534	92.9	72,580	97.10%
Lugar de residencia (País)	189,328	184,830	97.6	4,498	99.99%
Lugar de residencia (Estado)	6,103,439	5,624,943	92.2	478,496	98.52%
Lugar de residencia (Municipio)	6,103,439	5,624,943	92.2	478,496	96.17%
Lugar de nacimiento (País)	654,358	572,380	87.5	81,978	100.00%
Lugar de nacimiento (Estado)	20,729,826	20,077,835	96.9	651,991	99.98%
Lugar de asistencia escolar (País)	838	709	84.6	129	99.98%
Lugar de asistencia escolar (Estado)	444,654	415,418	93.4	29,236	97.76%
Lugar de asistencia escolar (Municipio)	444,654	415,418	93.4	29,236	97.18%
Lugar de origen (Estado)	3,570	3,225	90.3	345	99.16%
País destino (País)	11,337	10,753	94.8	584	99.98%
Causa de retorno	3,618	3,077	85.0	541	97.94%
Causa de migración	836,640	769,612	92.0	67,028	97.69%
Causa de emigración	4,225	3,361	79.6	864	95.75%

El porcentaje global de codificación automática fue del **96.6%**

Ocupación y actividad económica son las variables con el menor porcentaje de codificación automática, **69.9%** y **76.4%** respectivamente

Conociendo México

01 800 111 46 34

www.inegi.org.mx

atencion.usuarios@inegi.org.mx



INEGI Informa
