



**El servicio público
es de todos**

**Función
Pública**

Minería de datos para la identificación y clasificación de posibles conflictos de interés reportados

Julio de 2021

Contenido



Metodología



Desarrollo

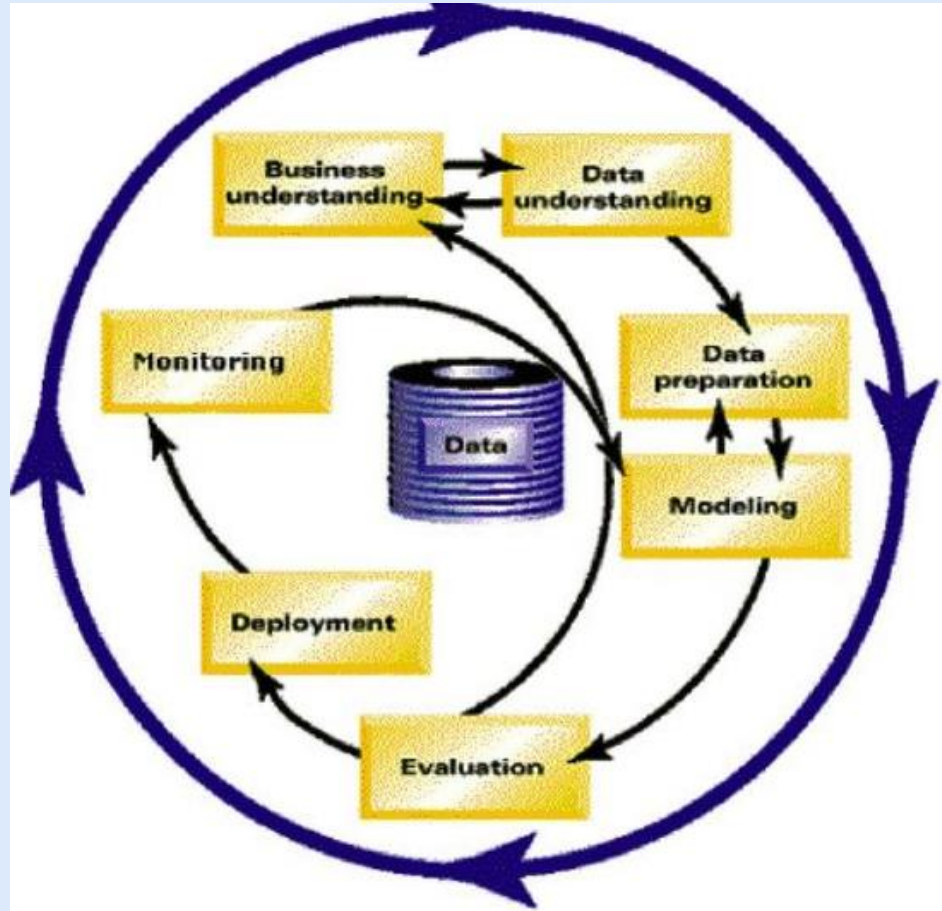


Conclusiones y Recomendaciones

1. Metodología



Fases de la metodología CRISP- DM



2. Desarrollo

```
scope.$watchWatchExpr, function ngSwitchMatchAction(value) {  
    // 1. 111  
    for (ii = 0, ii = previousElements.length; i < ii; ++i) {  
        previousElements[i].remove();  
    }  
    previousElements.length = 0;  
  
    for (ii = 0, ii = selectedScopes.length; i < ii; ++i) {  
        var selected = selectedElements[i];  
        selectedScopes[i].$destroy();  
        previousElements[i] = selected;  
        animate.leave(selected, function() {  
            previousElements.splice(i, 1);  
        });  
    }  
  
    selectedElements.length = 0;  
    selectedScopes.length = 0;  
  
    if ((selectedTranscludes = ngSwitchController.cases['!' + value] || ngSwitchC  
        scope.$eval(attr.change);  
        forEach(selectedTranscludes, function(selectedTransclude) {  
            var selectedScope = scope.$new();  
            selectedScopes.push(selectedScope);  
            selectedTransclude(selectedScope, function(clone, scope) {  
                clone.appendTo(scope);  
            });  
        });  
    }  
});
```

Entendimiento del Negocio

Identificación
de necesidad



Objetivos de
la minería de
datos



Alcance

Actualmente, según lo establece la Ley 2013 de 2019, Función Pública captura en el aplicativo dispuesto para ello, la información de declaraciones de bienes, rentas y conflictos de intereses de los sujetos obligados a declarar, sin embargo, esta información, en particular la correspondiente a los posibles conflictos de interés, no es de fácil interpretación y no está siendo procesada ni analizada, dado que esto demanda tiempo y recursos para realizarlo.

- ✓ Identificar y clasificar los posibles conflictos de interés registrados por los sujetos obligados.
- ✓ Parametrizar los posibles conflictos de interés registrados en el Aplicativo de Ley 2013 de 2019 para desarrollar productos y acciones en desarrollo de las políticas descritas.
- ✓ Analizar los perfiles y las características en común de los diferentes sujetos obligados que registran un posible conflicto de interés en el aplicativo.

- ✓ Para este ejercicio se aplicarán técnicas de minería de datos a la información de la base de enero a junio de 2021, correspondiente a aquellos sujetos obligados (persona natural y jurídica) que manifestaron un posible conflicto de interés en el formulario dispuesto para tal fin.

Análisis exploratorio descriptivo

Base de datos

- ✓ ID asociados hasta en 16 registros diferentes
- ✓ No hay capacidad de distinguir persona natural de persona jurídica
- ✓ 14 hojas distintas de información difícil emparejamiento.
- ✓ Excesiva dificultad para obtener resultados concluyentes.
- ✓ Demasiados campos nulos

Base de datos actualizada

- ✓ Se siguen repitiendo registros
- ✓ Permite distinción entre persona natural y jurídica
- ✓ Una sola hoja con toda la información
- ✓ Mayor facilidad para realizar análisis estadístico

Resultados

- ✓ 6 categorías son nulas en todos los registros.
- ✓ 149643 registros iniciales, de los cuales:
 - ☐ 121.006 tienen estado finalizado.
 - ☐ 101.132 son de ingreso.
 - ☐ 146.998 registrados con CC

1. Base personas naturales

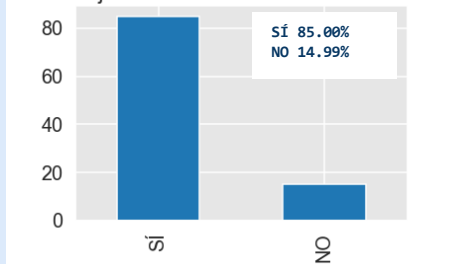
2. Base personas jurídicas

- ✓ Si se separan las bases, cada una tendría menos categorías, facilitaría análisis.

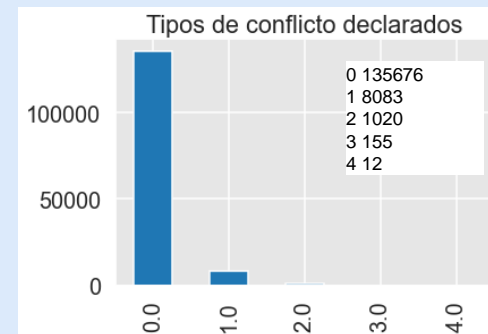
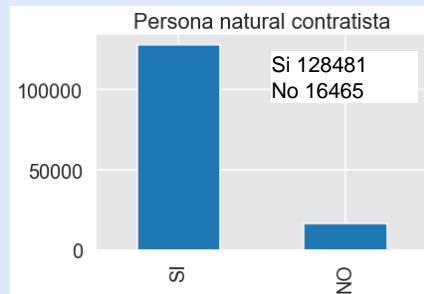
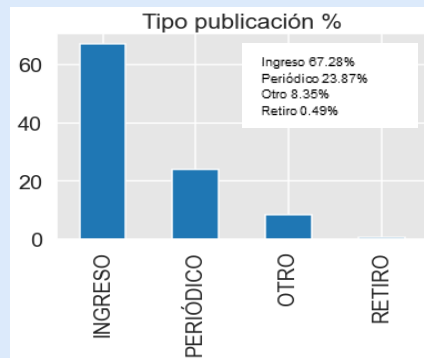
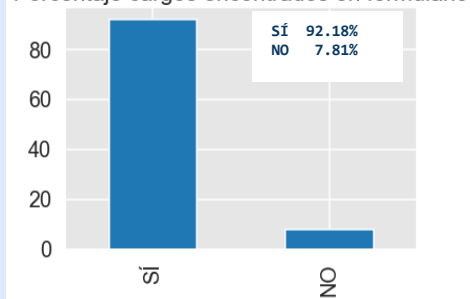
Análisis exploratorio descriptivo

Resultados persona natural

Porcentaje entidades encontradas en formulario

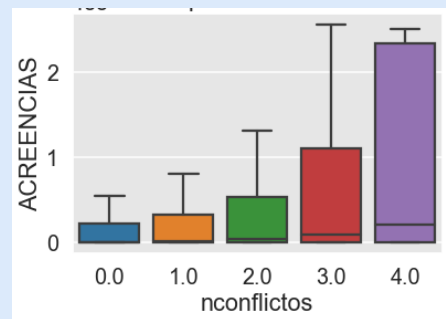
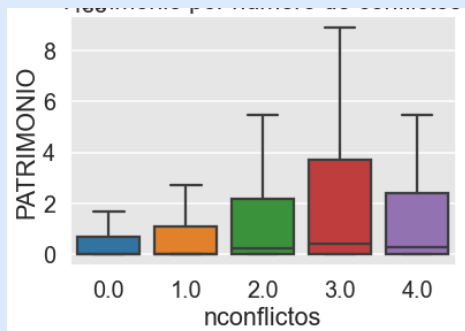
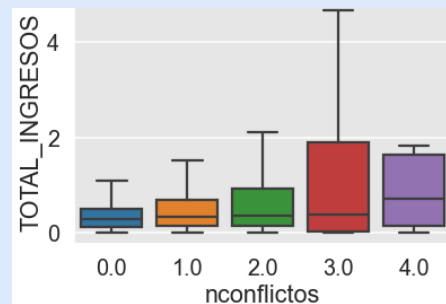
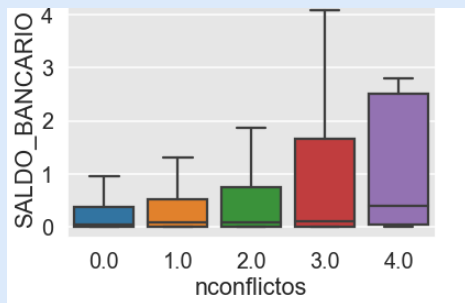


Porcentaje cargos encontrados en formulario



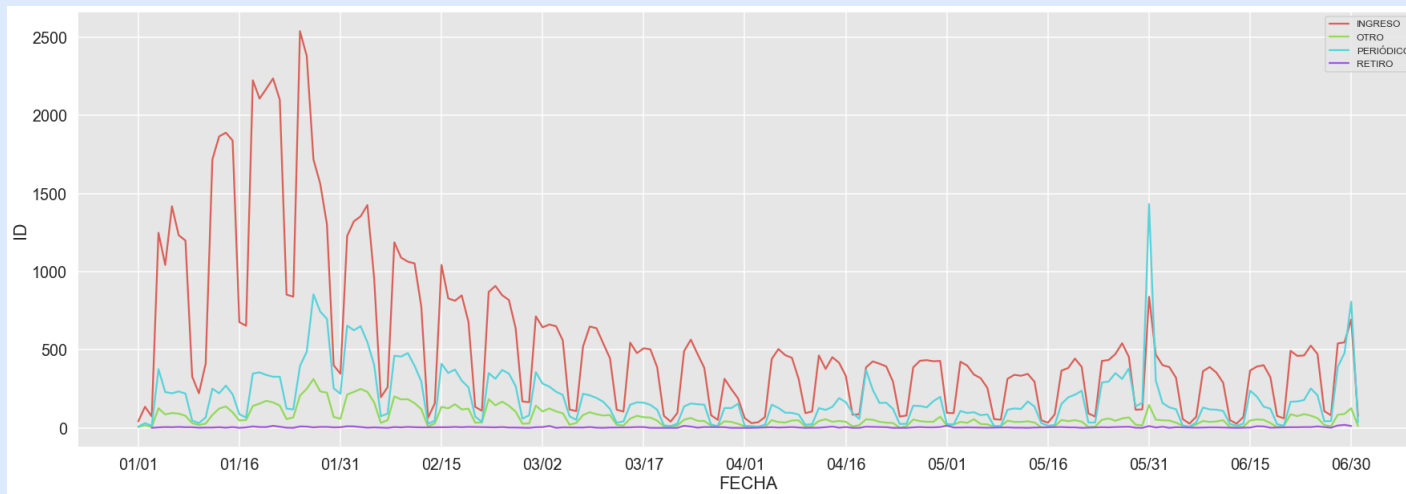
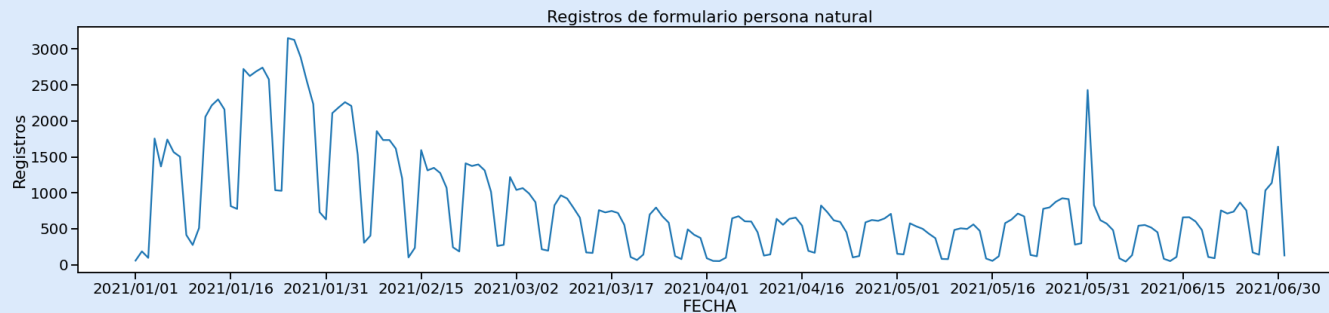
Análisis exploratorio descriptivo

Resultados persona natural



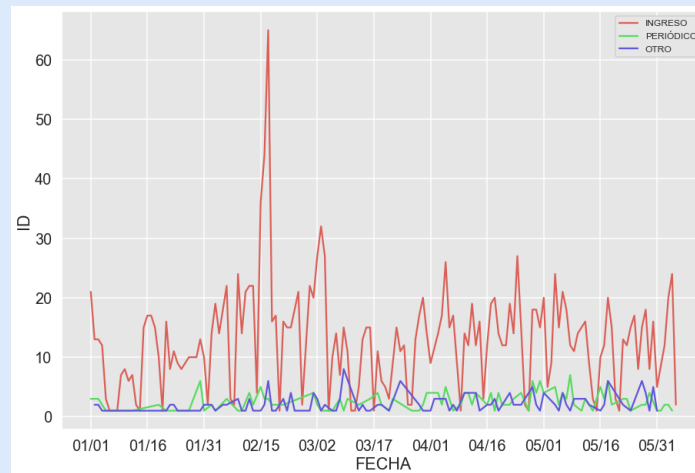
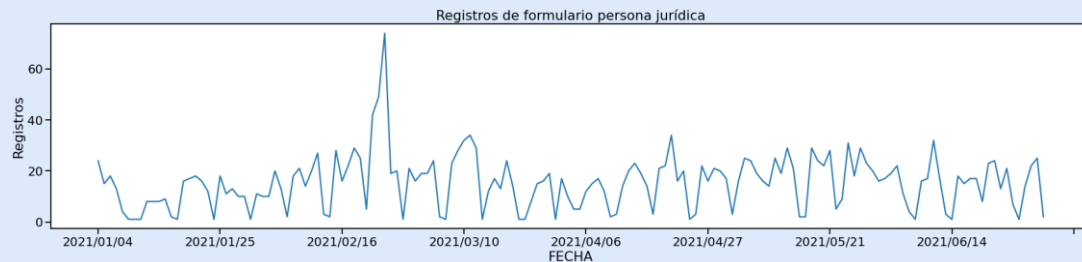
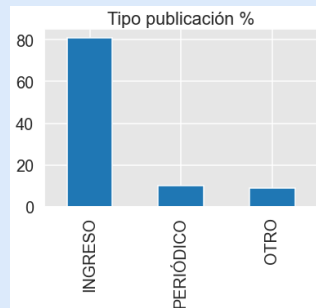
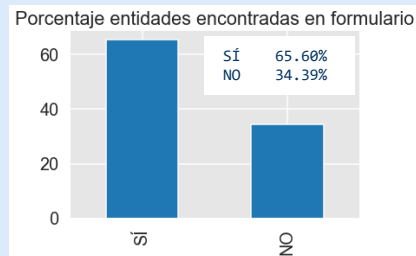
Análisis exploratorio descriptivo

Resultados persona natural



Análisis exploratorio descriptivo

Resultados persona jurídica



Entendimiento de los datos

Hoja en archivo Excel	Número total de registros	Registros usados para ejercicio	Descripción de problemas encontrados
CONYUGE_COMPANERO	133.406	0	La información disponible en la descripción no cumple con los criterios para minería de texto
DONACIONES	307	0	La información disponible en la descripción no cumple con los criterios para minería de texto
CI_PARIENTES	10.801	0	La información disponible en la descripción no cumple con los criterios para minería de texto
POSIBLES_CI	135.442	807	Se filtran los registros cuyo estado es finalizado o en revisión, además de los registros donde posible conflicto es "Sí". Por otro lado, existen casos donde en la descripción hay caracteres especiales o mencionan ninguno o no tengo, algunos de estos casos fueron excluidos.
CI_INFO_PJ	2.036	0	La información disponible en la descripción no cumple con los criterios para minería de texto

Procesamiento de datos

Enfoque de bolsa de palabras



- ✓ Convertir palabras a minúsculas
- ✓ Quitar espacios en blanco innecesarios
- ✓ Remover puntuación y caracteres especiales
- ✓ Remover stopwords
- ✓ Reducir las palabras a las raíces
- ✓ Calcular vector de características

Doc 1: casa bonita
Doc 2: niña bonita
Doc 3: casa grande

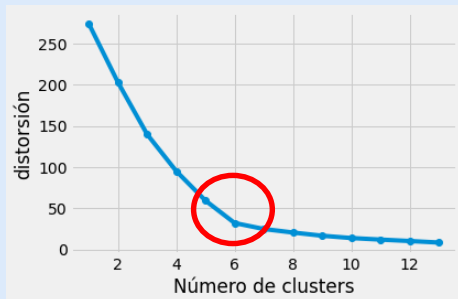


$Tf * idf$
 Tf

1	0	1
1	1	0
0	1	0
0	0	1



Técnicas supervisadas
Técnicas no supervisadas



Definición del número de grupos

Cluster	Número de registros	actividades agrícolas, comerciales, económicas	empresas vínculos	listas candidatos	sas empresas	sociedad
1	501	8,3266E-17	-2,4980E	-1,24903E	-1,3877E	1,2490E
2	78	0,0329	2,7755E	0,9877	-1,3877E	-1,3877E
3	65	0,9920	0,0146	-2,7755E-17	0,0048	0,0122
4	47	0,2772	0,0176	0,0272	0,0083	0,8938
5	68	0,0909	0,2316	0,0322	0,7948	0,209
6	48	0,0645	0,9321	0,0433	0,0257	0,0692

Procesamiento de datos

Clasificación	Descripción	Grupo	Total Registros
Organización, sociedad o asociación a la cual perteneció o continúa siendo miembro	Que el servidor sea socio de alguna de las personas interesadas en la actuación administrativa o su representante o apoderado en sociedad de personas.	4, 6	95
Lista de candidatos	Que el servidor haya hecho parte de listas de candidatos a cuerpos colegiados de elección popular inscritas o integradas también por el interesado en el período electoral coincidente con la actuación administrativa o en alguno de los dos períodos anteriores.	2	78
Relación contractual o de negocios	Que alguno de los interesados en la actuación administrativa sea representante, apoderado, dependiente, mandatario o administrador de los negocios del servidor público	3, 5	133
Otros	No fue posible categorizarlos por las descripciones proporcionadas por el usuario.	1	501

Análisis de texto entidades nombradas

Tipo	Entidad
Relación con las partes	Misma entidad, apoderado, convenio, contratos, contratación, asociación, representante legal.
Organización, sociedad o asociación a la cual perteneció o continúa siendo miembro	Organización, sociedad, asociación, actividad de negocio, tienda, local, negocio, almacén, emprendimiento , establecimiento.
Acreedor/deudor	Deudor, crédito, préstamo, hipoteca, acreedor.
Heredero o legatario	Herencia, heredero, sucesor.

Tipo	Entidad
Litigio o controversia/ decisión administrativa pendiente	Litigio, litigante, controversia, denuncia, pleito, decisión administrativa, demanda.
Denuncia penal o disciplinaria	Denuncia, penal , disciplinario, actuación administrativa.
Curador o tutor del interesado	Curador, tutor.
Participación directa/asesoría de alguna de las partes interesadas	Asesor, socio.

Identificación y clasificación de conflictos

Posibles conflictos de interés por cónyuge o compañero

Total Registros	1.783
Sin clasificar*	1.518
Posibles conflictos	215

*Descripción	Total
Ninguno	423
Esposo, unión libre, cónyuge, compañero	218
Compartimos gastos, abogado, casa	127
Otros	750



Identificación y clasificación de conflictos

Posibles conflictos de interés por representante legal

Total Registros	56
Sin clasificar*	44
Posibles conflictos	12

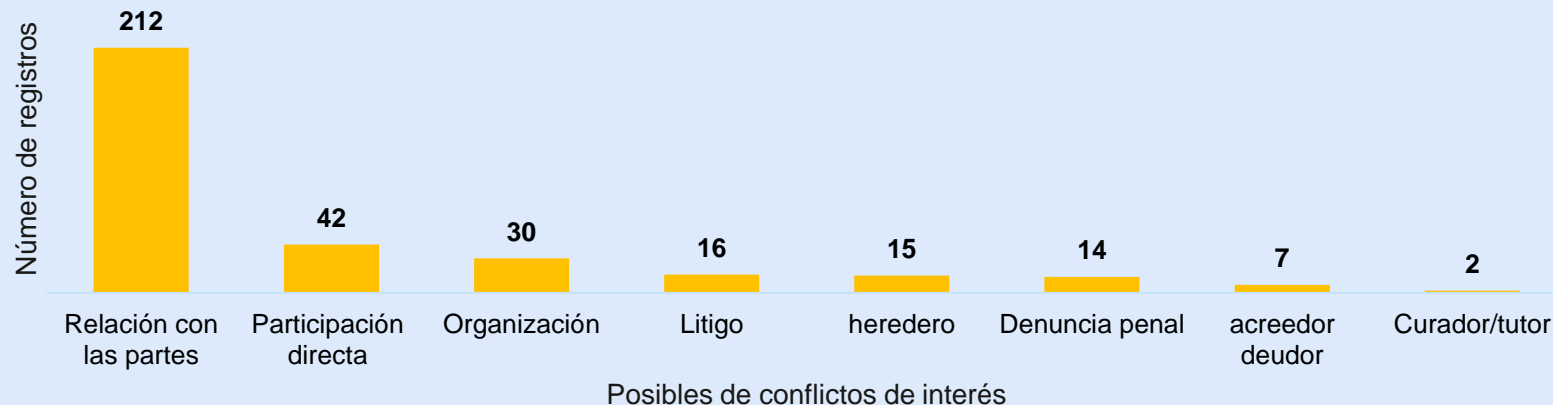


*32 sujetos obligados en la descripción manifiestan no tener ningún conflicto de interés

Identificación y clasificación de conflictos

Posibles conflictos de interés por parientes

Total Registros	5.778
Sin clasificar*	4.593
Posibles conflictos	338



*695 sujetos obligados en la descripción manifiestan no tener ningún conflicto de interés

Identificación y clasificación de conflictos

Posibles conflictos de interés adicionales

Total Registros	1.107
Sin clasificar*	559
Posibles conflictos	554



*244 sujetos obligados en la descripción manifiestan no tener ningún conflicto de interés

Conclusiones



Conclusiones

- ✓ En total se identificaron y clasificaron 1.113 posibles conflictos de interés

Conflicto de interés	Cantidad	Participación
Relación con las partes	441	40%
Organización	273	25%
Participación directa	143	13%
Litigo	114	10%
Lista de candidatos	82	7%
Otros	60	5%
Total	1.113	100%

- ✓ De los 1.113 conflictos de interés clasificados el 62% corresponde a sujetos obligados que manifiestan ser contratistas.
- ✓ En la sección de marcación de conflicto de interés por litigio controversia o denuncia penal a pesar de que 43 sujetos obligados marcaron que si existe un posible conflicto solo existen dos descripciones relacionadas con el tema, en cuanto a los 41 restantes la descripción es que no existe ningún conflicto de interés.
- ✓ En la sección de conflicto de interés por amistad o enemistad se encontraron 45 sujetos obligados que manifestar presentar un posible conflicto de interés por este tipo, sin embargo, las descripciones no corresponden a un posible conflicto de interés o son que no existe ningún conflicto.
- ✓ El enfoque de entidades nombradas tiene un mejor resultado para los datos disponibles en el aplicativo con una clasificación del total de registros de un 49% superior al 38% del ejercicio realizado con el enfoque de bolsa de palabras.

Recomendaciones

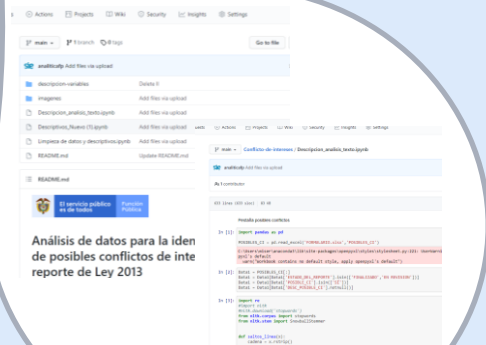
- ✓ Parametrizar el sistema de tal forma que no permita marcar un si y en la descripción incluir signos o símbolos o que no se tiene ningún conflicto de interés
- ✓ Reforzar capacitaciones con los sujetos obligados para el correcto diligenciamiento del aplicativo
- ✓ Establecer algún tipo de lista desplegable de conflictos de interés de acuerdo con la normatividad, con la cual los sujetos obligados se puedan identificar y de esta manera se facilita la clasificación de los posibles conflictos de interés
- ✓ Analizar posibles cambios metodológicos que propendan por impedir la multiplicidad de registros, ya que esto afecta la calidad de los datos para futuros ejercicios de análisis
- ✓ Llevar un registro central que permita resumir los principales resultados de las declaraciones de bienes y rentas, se debe pensar en una base de datos en un solo conjunto que permita simultáneamente visualizar los diferentes campos marcados por cada individuo.
- ✓ El registro de personas naturales y jurídicas debería ser independiente uno de otro, sin embargo ante la dificultad que esto pueda generar, se sugiere mantener un índice numérico que permita en todas las bases de datos distinguir ambos tipos de declarantes.

Publicación de los resultados



Repositorio público o privado que contiene todos los archivos de un proyecto, así como el historial de revisión de cada archivo. Los repositorios pueden tener varios colaboradores y pueden ser públicos o privados.

<https://github.com/analiticafp/Conflicto-de-intereses.git>





El servicio público
es de todos

Función
Pública

Gracias

Carrera 6 No 12-62, Bogotá D.C., Colombia



7395656 Fax: 7395657



Línea gratuita de atención al usuario: 018000 917770



www.funcionpublica.gov.co



eva@funcionpublica.gov.co