

Dirección de Desarrollo Digital

Unidad de Científicos de
Datos



**El futuro
es de todos**

DNP
Departamento
Nacional de Planeación



CARACTERIZACIÓN Y COMPARACIÓN REGIONAL DE TRÁMITES EMPRESARIALES Y DE ALTO IMPACTO

INFORME FINAL

Dependencias y entidades involucradas	Departamento Nacional de Planeación <ul style="list-style-type: none">• Dirección de Desarrollo Digital - Unidad de Científicos de Datos• Dirección de Innovación y Desarrollo Empresarial – Subdirección de Productividad, Internacionalización y Competencia.• Función Pública.• Ministerio de Tecnologías de la Información y las Comunicaciones.
Sector	Ciencia, tecnología e innovación
Tecnologías utilizadas	Python
Fuentes de datos	<ul style="list-style-type: none">• Información registrada en el SUIT de los trámites u otros procedimientos administrativos (ver)• Datos de operación de trámites y OPAs (ver)• Interacción likes/dislikes de trámites en gov.co• Base de datos de audiencia de trámites

Contenido

1. Presentación	2
2. Objetivos del proyecto	2
3. Metodología	2
4. Resultados	6
5. Conclusiones y recomendaciones	12
6. Socialización	13
Contacto	13



1. Presentación

En el marco del CONPES 3956 “Política de Formalización Empresarial” se evidencia la importancia de mejorar la relación beneficio costo de ser formal. Un aspecto clave que contribuye a este objetivo es el de simplificar, interoperar y optimizar los trámites que afectan a los empresarios y en general a la ciudadanía. Por consiguiente, en conjunto con el Ministerio de Tecnologías de la Información, Función Pública y el Departamento Nacional de Planeación (DNP); se desarrolla este proyecto que consiste en caracterizar y comparar los trámites ofrecidos tanto en entidades nacionales como territoriales a partir de la información contenida en las bases de datos del SUIT y de GOV.CO. Esto con el fin de identificar prioridades de interoperabilidad y acciones de optimización. Para ello, se desarrolló una metodología basada en procesamiento de texto para comparar las descripciones de los trámites y encuentra semejanzas entre ellos. Adicional se construyó una herramienta de visualización y comparación de trámites empresariales par facilitar la tarea de analisis.

Within the framework of CONPES 3956 “Business Formalization Policy”, the importance of improving the cost-benefit ratio of being formal is evident. A key aspect that contributes to this objective is to simplify, interoperate and optimize the procedures that affect employers and citizens in general. Consequently, in conjunction with the Ministry of Information Technology, Public Function, and the National Planning Department (DNP); This project is being developed, which consists of characterizing and comparing the procedures offered both in national and territorial entities from the information contained in the SUIT and GOV.CO databases, to identify interoperability priorities and optimization actions. To do this, a methodology based on word processing was developed to compare the descriptions of the procedures and find similarities between them. Additionally, a visualization and comparison tool for business procedures was built to facilitate the analysis task.

2. Objetivos del proyecto

2.1. General

Caracterizar los trámites ofrecidos por las entidades nacionales y territoriales, a partir de la información contenida en las 4 bases de datos mencionadas anteriormente, para identificar similitudes entres trámites, con el fin de identificar opciones de simplificación de trámites, prioridades de interoperabilidad y estimación de trámites dirigidos a empresas.

2.2. Específicos

1. Realizar análisis descriptivo de las cuatro bases de datos para caracterizar los trámites que están dirigidos a micros, pequeñas, medianas y grandes empresas.
2. Implementar rutina de *web-scraping*¹ para complementar información adicional de los trámites reportados en el SUIT tales como su duración, documentos, y el propósito del trámite.
3. Desarrollar metodología de analisis y comparación de texto para identificar similitudes entre trámites a partir de su propósito, con el objetivo de encontrar opciones de simplificación de trámites empresariales y prioridades de interoperabilidad.
4. Desarrollar herramienta de visualización para facilitar la visualización de operaciones y la comparación de trámites empresariales.

3. Metodología

La metodología de este proyecto se divide en 3 fases: (i) el analisis descriptivo sobre las 4 bases de datos presentadas para el proyecto, (ii) La obtención de información adicional para los trámites de interés a partir de técnicas de *web-scraping* y (iii) la metodología para comparar trámites a partir de su propósito (descripción) de cada trámite. Cada una de estas fases se explica a continuación.

3.1. Analisis descriptivo de los datos

Para el desarrollo del proyecto se contó con cuatro (4) bases de datos.

¹ Web scraping o raspado web, es una técnica utilizada mediante programas de software como Python para extraer información de sitios web.



- Información registrada en el SUIT de los trámites u otros procedimientos administrativos: contiene la información registrada en el Sistema Único de Información de Trámites (SUIT) por las entidades a nivel nacional y territorial. Allí se encuentran los detalles de cada uno de los trámites, procedimientos administrativos y consulta de información, tales como la entidad prestadora, número único, nombre, propósito, departamento, municipio, entre otros datos propios de cada trámite. Esta base de datos tiene 510,263 registros correspondientes a 65,888 trámites registrados en SUIT.
- Datos de operación de trámites y OPAs: contiene la información de operaciones o transacciones por cada trámite que han sido resueltos por la entidad que lo presta. En esta base de datos se encuentra información de la entidad, número único, número de solicitudes resueltas (de manera presencial, virtual o parcial) por año (desde 2014), número de PQRDS resueltas entre otras. Esta base de datos contiene 45,208 registros de operaciones de trámites, que corresponden a 16,559 trámites.
- Interacción likes/dislikes de trámites en gov.co: base de datos que incluye para algunos trámites las interacciones que hacen los usuarios y su percepción acerca de ellos. La percepción es expresada mediante una reacción de gusto, disgusto o solicitud de mejora del trámite, además, se incluyen los comentarios que los usuarios de los trámites dejan al realizarlos. La información de interacción está disponible para los años 2020 y 2021 con corte a marzo del 2021. Esta base contiene 30,469 registros, los cuales corresponden a 3,016 trámites.
- Base de datos de audiencia de trámites: contiene la información de la identificación de trámites que van dirigidos a micro, pequeña, medianas o grandes empresas. Esta base de datos tiene 71,978 de registros de trámites, de los cuales 274 están dirigidos a empresas y han sido verificados por la entidad que presta el servicio.

El objetivo principal de esta etapa fue entender, consolidar e identificar las variables que ayudan a responder la necesidad de prioridades de interoperación de los trámites dirigidos a empresas a partir de la percepción de dificultad de realizar cierto trámite, estimado a partir del número de PQRSD recibidas o impresiones de disgusto del trámite. Los resultados de esta fase se consignan en la Sección 4.1.

3.2. Web-scraping para obtener información adicional

Las técnicas de *web-scraping* permiten obtener información de internet mediante navegación automática en un portal definido. En el caso específico, se automatizó el proceso de búsqueda información adicional para los 65,888 trámites que están registrados en el [SUIT](#). El primer paso fue identificar el patrón de la url para la búsqueda de información adicional, en la Figura 1, se ilustra que para llegar a la hoja de vida del trámite (ver Figura 2) solo se necesita su identificador único.

Figura 1: Estructura de la URL para la búsqueda automatizada de trámites.



URL estandar del visor SUIT

Identificador del trámite

Fuente: Elaboración propia.

De cada hoja de vida del trámite (ver Figura 2) se extrae de forma automática información que no está reportada en ninguna de las 4 bases de datos del proyecto o campos que se encuentra incompletos:

- *¿Requiere pago?* (No está en las bases de datos)
- *¿Es totalmente en línea?* (No está en las bases de datos)
- *Descripción.* (Información incompleta dentro de las bases de datos)
- *2- Alistar documentos.* (Información incompleta dentro de las bases de datos)
- *Texto del resultado que se obtiene.* (Información incompleta dentro de las bases de datos)
- *Texto del tiempo en el que se obtiene el resultado.* (No está en las bases de datos)



Como resultado de esta fase se obtiene la información completa para cada trámite. De aquí, la información de la descripción del trámite será la que se usará en la etapa de calculo de similitudes. La información adicional restante es usada para el tablero de visualización de comparación de trámites empresariales.

Figura 2: Formulario en línea de la hoja de vida de cada trámite del portal SUIT. En color lila, la información que se extrae mediante web-scraping para cada trámite.

<p>Información proporcionada por:</p> <p>UNIDAD ADMINISTRATIVA ESPECIAL DIRECCIÓN DE IMPUESTOS Y ADUANAS NACIONALES</p> 	Actualización del Registro Único Tributario (RUT)	
	<i>(También se conoce como: Actualizar el RUT)</i>	
	¿Cuándo se puede realizar?	Cualquier fecha
	¿A dónde ir?	Ver puntos de atención
	¿Requiere pago?	No, es gratuito
	¿Es totalmente en línea?	No
	Última actualización: 28 Julio 2021	
Descripción		
<p>Con este trámite obtendrá: Realizar modificaciones en la información registrada en el Registro Único Tributario (RUT). ¿Cuándo debo actualizar el RUT? • Cuando vaya actualizar datos personales (identificación y ubicación). • Datos comerciales (actividad económica, responsabilidades, entre otros). • Por retiro de responsabilidad del IVA o de otras obligaciones tributarias. • Si tiene un cambio de representación legal o apoderado. ¿Qué pasa si no actualizo el RUT? De acuerdo con el Estatuto Tributario (art. 658-3), se podría aplicar una multa de entre 1 y 2 Unidad de Valor Tributario (UVT) (\$ 36.308 para 2020), por cada día de retraso en la actualización de la información.</p>		
Para realizarlo necesita:		
<div style="display: flex; justify-content: space-between; border-bottom: 1px solid #ccc; padding-bottom: 5px;"> Ciudadano Extranjeros Instituciones o dependencias públicas Organizaciones </div> <ol style="list-style-type: none"> <li style="border: 1px solid #ccc; padding: 5px; margin-bottom: 5px;">1 - Solicitar cita + <li style="border: 2px dashed purple; border-bottom: 1px solid #ccc; padding: 5px; margin-bottom: 5px;">2 - Alistar documentos y cumplir requisitos + <li style="border: 1px solid #ccc; padding: 5px; margin-bottom: 5px;">3 - Actualizar RUT en el portal DIAN + <li style="border: 1px solid #ccc; padding: 5px; margin-bottom: 5px;">4 - Radicar documentos + <li style="border: 1px solid #ccc; padding: 5px; margin-bottom: 5px;">5 - La actualización se verá reflejada en el RUT al momento de acudir a su cita o al hacerlo de manera virtual + 		
Resultado que se obtiene		
<p>★ Formulario actualizado del Registro Único Tributario (RUT) que se obtiene en 20 Minuto(s)</p>		

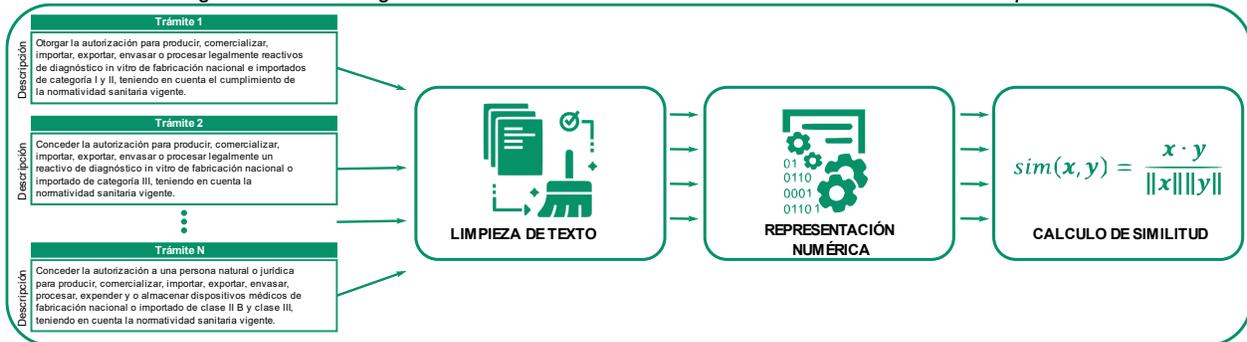
Fuente: Elaboración propia.

3.3. Comparación de trámites a partir de su descripción

Con el objetivo de identificar trámites similares entre sí, se propuso una metodología basada en análisis de texto para identificar similitudes a partir de las descripciones de cada trámite. En la Figura 3 se ilustra los pasos para calcular la similitud entre trámites. Primero, los textos de las descripciones pasan por un proceso de limpieza de texto. Luego cada texto se convierte a una representación numérica. Por último, estas representaciones se comparan una contra todas mediante la similitud coseno, obteniendo de esa forma pares de similitudes entre trámites. A continuación, se explica todo el proceso en detalle.



Figura 3: Metodología del cálculo de similitud de trámites basados en su descripción.



Fuente: Elaboración propia.

• Limpieza de texto

Para esto, cada descripción pasa por una limpieza de texto de 4 pasos:

1. Remover signos de puntuación y conversión a minúsculas: se retiran todos los signos de puntuación del texto tales como comas, puntos, signos de exclamación. Después todo el texto es convertido a texto en letras minúsculas.
2. Remoción de *stopwords* o palabras vacías: estas palabras no tienen significado por sí solas, sino que modifican o acompañan a otras. Este conjunto se encuentra conformado por artículos, pronombres, preposiciones y adverbios.
3. Quitar palabras propias del contexto: Se remueve del texto palabras que pueden ser muy frecuentes dado la naturaleza del estudio, para ello se removieron nombres propios de personas, municipios y departamentos de Colombia y números.
4. Lematización de texto: se refiere a la transformación de todas las formas flexionadas de una palabra en su lema, por ejemplo, canto, cantas, cantamos, cantan son distintas formas (conjugaciones) del verbo cantar, por lo que todas estas palabras son representadas por este lema.

En la Figura 4, se muestra la descripción del trámite con identificador único 33529, después del proceso de limpieza de texto. Este proceso ayuda a estandarizar los textos de las descripciones y facilita el proceso de comparación entre ellas para encontrar trámites similares entre sí.

Figura 4: Ejemplo del resultado de la limpieza de texto



Fuente: Elaboración propia.

• Similitud entre descripciones de trámites

Una vez se ha realizado la limpieza de los textos, estos se deben transformar a una representación numérica que facilite el proceso de cuantificar que tan similares son dos descripciones. Para esto se utiliza la metodología TF-IDF (term frequency-inverse document frequency). TF-IDF es una medida estadística utilizada para evaluar la importancia de una palabra para una descripción en un conjunto de descripciones. Es decir, la importancia de la palabra aumenta proporcionalmente al número de veces que aparece en la descripción (parte TF), pero se compensa con la frecuencia



de la palabra en las otras descripciones que hacen parte del conjunto de tramites analizados (parte IDF). Matemáticamente se representa de la siguiente manera:

$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right)$$

Donde $w_{i,j}$ es la importancia de la palabra i en la descripción j , $tf_{i,j}$ es el número de ocurrencias de la palabra i en la descripción j , df_i es el número de descripciones que contienen a i , y N es el total de descripciones de trámites que hacen parte del análisis. Entonces, si la palabra es muy común y aparece en muchas descripciones, este número se acercará a 0. De lo contrario, se acercará a 1.

A partir de esto cada descripción se transforma en un vector numérico de tamaño fijo que puede ser comparado con otro a partir de medidas de distancia o similitud entre vectores. Estos vectores se construyen con la ubicación de cada palabra en un diccionario de términos y su importancia calculada con la metodología TF-IDF. En la Figura 5, se muestra una representación numérica construida con un diccionario de 6 términos (actualizar, distribuir, importar, obtener, permiso y resultado) para 3 descripciones de ejemplo.

Figura 5: Ejemplo de vectores numéricos construidos con la metodología TF-IDF

	actualizar	distribuir	importar	obtener	permiso	resultado
descr_1	0.234277	0.304905	0.570708	0.799959	0.211379	0.206535
descr_2	0.310522	0.713638	0.227315	0.437324	0.884007	0.929069
descr_3	0.298243	0.485664	0.374303	0.198483	0.354976	0.934600

Fuente: Elaboración propia.

Después, para cuantificar la similitud entre descripciones de trámites se calcula la medida de similitud coseno entre cada par de representaciones numéricas (vectores numéricos) de la base de trámites. Esta medida de similitud se expresa matemáticamente:

$$sim(x, y) = \frac{x \cdot y}{\|x\| \|y\|}$$

Donde x y y son los vectores para comparar, el operador $\|\cdot\|$ representa la norma euclidiana del vector, que conceptualmente es la longitud de cada vector. Un valor cercano a 0 significa que no hay similitud entre los vectores. Un valor cercano a 1 significa una alta similitud entre dos vectores.

A partir de la similitud coseno, se puede cuantificar si dos tramites son similares tomando la representación numérica de las descripciones o propósitos cada trámite. Los resultados de la metodología de comparación de trámites se presentan en la Sección 4.2.

4. Resultados

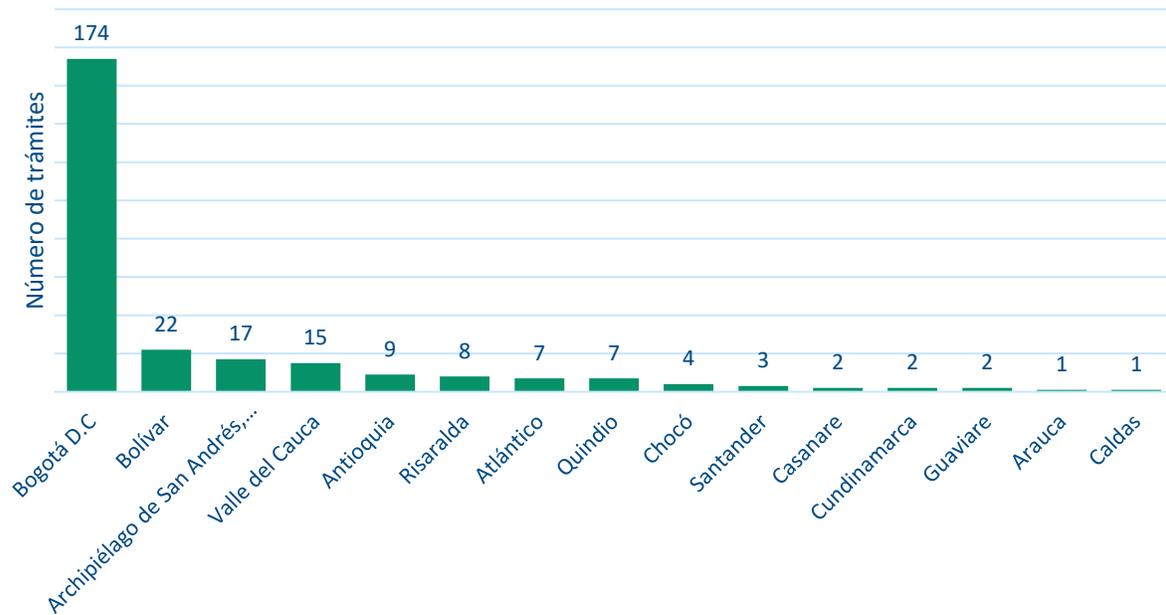
A través del desarrollo metodológico descrito en la Sección 3, se obtuvieron los resultados que se presentan a continuación.

4.1. Resultados del analisis descriptivo de las bases de datos

Dado que el enfoque del proyecto está pensado para trámites empresariales, se cruzó la base de datos “audiencia de trámites” con “base de registro de trámites del SUIT”. Con esto se logró tener la información al detalle de los 274 trámites empresariales. De este cruce se identifican los trámites empresariales disponibles por región, en la Figura 6, se observa que solo 15 regiones han categorizado trámites dirigidos a empresas; Además, el 64% está ubicado en Bogotá.



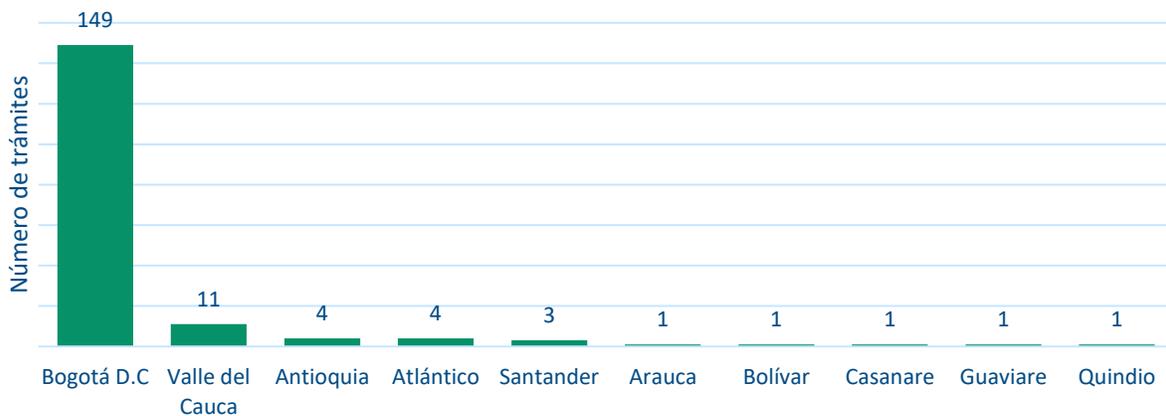
Figura 6: Trámites empresariales reportados por las entidades por territorio.



Fuente: Elaboración propia.

Esta información previa, se complementa con la información sobre el registro de operaciones de los trámites reportada desde el año 2014 en adelante. Se identifica que, de los 274 trámites empresariales, 176 reportan operación, es decir, al menos se resolvió una solicitud por la entidad prestadora del trámite. Esto reduce a 10 territorios en los que se ha registrado al menos una operación de trámites dirigidos a empresas. Esto se ilustra en la Figura 7.

Figura 7: Número de trámites empresariales por territorio que registraron al menos una operación desde el 2014.



Fuente: Elaboración propia.

A nivel de entidad prestadora de trámites dirigidos a empresas, 60 de ellas han identificado al menos uno. Estas 60 entidades corresponden al 2,4% del total de entidades registradas en la base SUIT. En la Figura 8, se muestra el top 10 de entidades que han identificado trámites dirigidos a empresas. La DIAN y la Superintendencia de Vigilancia y Seguridad Privada son las que más han identificado trámites empresariales.



Figura 8: Número de trámites empresariales por entidad que registraron al menos una operación desde el 2014.



Fuente: Elaboración propia.

Por otro lado, la información de número de PQRDS recibidas por cada trámite empresarial puede ser útil para orientar la búsqueda de simplificación de trámites, bajo es supuesto que, entre mayor el número de PQRDS, mayor dificultad para obtener el resultado final del trámite. En la Tabla 1, se muestra el top 15 de los trámites empresariales que han recibido mayor numero de PQRDS desde el 2014 a la actualidad. La Tabla 1 también muestra que los trámites relacionados con la inscripción, actualización y cancelación del RUT son los que mayor PQRDS reciben. Esto puede verse como uno de los trámites que presenta mayores problemas para obtener un resultado satisfactorio.

Tabla 1: Número de PQRDS recibidos por trámite desde el 2014.

Nombre del trámite	PQRDS
Actualización del Registro Único Tributario (RUT)	87.394
Certificado de Carencia de Informes por Tráfico de Estupefacientes	47.074
Inscripción en el Registro Único Tributario (RUT)	26.762
Cancelación de la inscripción en el Registro Único Tributario (RUT)	26.148
Afiliación al Fondo Nacional de Ahorro (Cesantías/Ahorro Voluntario Contractual).	21.080
Levantamiento de la suspensión de la inscripción en el Registro Único Tributario (RUT)	13.662
Facilidades de pago cobro contribuciones parafiscales	7.731
Pago seguro de depósitos	7.466
Registro de TIC - Industria de Comunicaciones	5.973
Inscripción de establecimientos, entidad e institución que produzca o provea bienes y servicios normados como objeto de vigilancia y control sanitario.	5.196
Autorizaciones extraordinarias para el manejo de sustancias químicas controladas	3.055
Registro sanitario, permiso sanitario, notificación sanitaria para alimentos.	2.530
Autorización de venta de equipos terminales móviles	2.058
Registro sanitario o renovación de medicamentos importados incluidos en normas farmacológicas colombianas	2.020
Registro sanitario de medicamentos de fabricación nacional nuevos y/o renovaciones incluidos en normas farmacológicas colombianas	1.770

Fuente: Elaboración propia.

En la Tabla 2 se muestra el top 15 de los trámites empresariales que han recibido mayor número de PQRDS para los dos últimos años, se observa que el comportamiento se mantiene con respecto a los trámites relacionados con el RUT. También se observa que el trámite “Certificado de Carencia de Informes por Tráfico de Estupefacientes” ocupa el segundo lugar en los dos rankings.



Tabla 2: Número de PQRDS recibidos por trámite para los años 2020 y 2021.

Nombre del trámite	PQRDS
Actualización del Registro Único Tributario (RUT)	18.182
Certificado de Carencia de Informes por Tráfico de Estupefacientes	14.222
Cancelación de la inscripción en el Registro Único Tributario (RUT)	5.872
Inscripción en el Registro Único Tributario (RUT)	5.664
Facilidades de pago cobro contribuciones parafiscales	4.857
Autorizaciones extraordinarias para el manejo de sustancias químicas controladas	3.055
Registro de TIC - Industria de Comunicaciones	1.551
Certificación para la exclusión del Impuesto al Valor Agregado (IVA)	1.522
Levantamiento de la suspensión de la inscripción en el Registro Único Tributario (RUT)	822
Registro sanitario, permiso sanitario, notificación sanitaria para alimentos.	775
Registro sanitario o renovación de medicamentos importados incluidos en normas farmacológicas colombianas	760
Registro sanitario de medicamentos de fabricación nacional nuevos y/o renovaciones incluidos en normas farmacológicas colombianas	681
Afiliación al Fondo Nacional de Ahorro (Cesantías/Ahorro Voluntario Contractual).	600
Autorización de venta de equipos terminales móviles	377
Inscripción de establecimientos, entidad e institución que produzca o provea bienes y servicios normados como objeto de vigilancia y control sanitario.	272

Fuente: Elaboración propia.

Partiendo de la información registrada en la base de datos de interacciones, se toma la información de dislikes como una percepción negativa del trámite que denota dificultad para realizar el mismo. En la Tabla 3 se muestra el top 10 de los trámites con mayor cantidad de dislikes. Se observa al igual que las tablas anteriores que los trámites relacionados con el RUT son los que tienen más percepción negativa. Además 6 de los 10 trámites reportados son prestados por la DIAN.

Tabla 3: Número de dislikes (no me gusta) reportados para trámites empresariales.

Nombre del trámite	Dislikes
Inscripción en el Registro Único Tributario (RUT)	2.498
Actualización del Registro Único Tributario (RUT)	1.912
Cancelación de la inscripción en el Registro Único Tributario (RUT)	127
Estado de cuenta de aportes parafiscales	10
Registro de libros de contabilidad	9
Asignación, reconocimiento o renovación de código de notificación sanitaria obligatoria para productos de higiene doméstica	8
Levantamiento de la suspensión de la inscripción en el Registro Único Tributario (RUT)	6
Registro sanitario o renovación de medicamentos importados incluidos en normas farmacológicas colombianas	5
Certificado de Carencia de Informes por Tráfico de Estupefacientes	4
Inscripción, renovación, ampliación, modificación o cancelación para realizar cualquier tipo de actividad relacionada con el manejo de materias primas de control especial y/o medicamentos que las contengan FNE	3

Fuente: Elaboración propia.

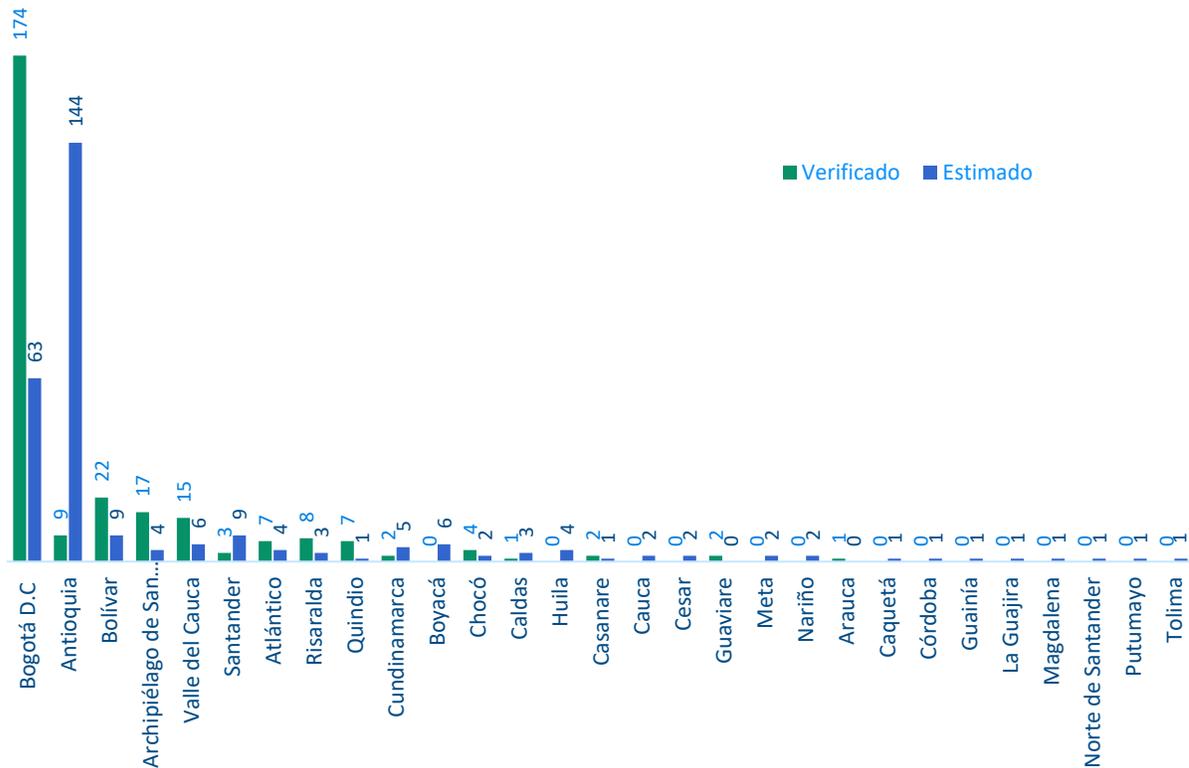


4.2. Resultados de comparación de trámites

Cada una de las descripciones (o propósito del trámite) de los 65.888 trámites presentes en la base del [SUIT](#)², pasan por las etapas descritas en la Sección 3.3. A partir de ahí, se calcula cada par de similitudes de trámites, dando como resultado una matriz triangular inferior de 65.897×65.897 , donde el valor de la posición i,j representa el grado de similitud que hay entre esos dos trámites (ver ANEXO 2). Dada la gran cantidad de información, esta gran matriz es filtrada a solo los trámites verificados como empresariales (274) por las entidades y los tramites que presentan una similitud mayor al 60% contra los tramites verificados. Bajo ese límite de similitud (60%), se identifican 280 trámites adicionales, es decir, que pueden ser dirigidos a empresas por su alta similitud con los trámites verificados como empresariales. Esto da como resultado una matriz triangular inferior de 554×554 .

En la Figura 9, se observa el número de trámites estimados basado en la similitud con el conjunto de trámites empresariales por territorio. También se puede observar que el número de territorios que ofrecen trámites empresariales aumentó de 15 a 30 territorios.

Figura 9: Número de trámites empresariales verificados y estimados por territorio.



Fuente: Elaboración propia.

A nivel de número de entidades que prestan trámites empresariales pasó de 60 a 284 cuando se incluyen los trámites identificados a partir de la similitud de descripciones. En la Figura 10, se muestra el top 10 de las entidades que tuvieron mas trámites empresariales identificados basados en la similitud de texto.

² Detalle de la información registrada en el SUIT por las entidades a nivel nacional y territorial de los trámites u otros procedimientos administrativos y consultas de acceso a la información pública.



Figura 10: Top 10 de las entidades con más trámites empresariales identificados basados en la similitud entre descripciones.



Fuente: Elaboración propia.

4.3. Herramienta para la visualización de operaciones y la comparación de trámites empresariales.

Para facilitar la visualización y comparación de trámites dirigidos a empresas y con el objetivo de dar insumos para la simplificación de trámites y la priorización de interoperabilidades, se desarrolló una herramienta de visualización que permite tres cosas principalmente:

1. Monitorear las solicitudes y número de PQRDS que reciben los trámites identificados como empresariales que reportan las entidades territoriales y nacionales en la base de datos "Datos de operación de trámites y OPAs" ([ver](#)), en el portal de datos.gov.co.
2. Ver información detallada de los trámites empresariales, tales como la entidad prestadora, resultado y tiempo en el que se obtiene el trámite, documentos que se requieren para solicitar el trámite, entre otros. Además, permite hacer paralelos de información entre dos trámites para ver las características comunes y diferentes entre ellos. La Figura 11 ilustra un ejemplo de comparación entre el trámite 2302 y el trámite 2335.

Figura 11: Panel de comparación de información entre trámites dentro de la herramienta de visualización.

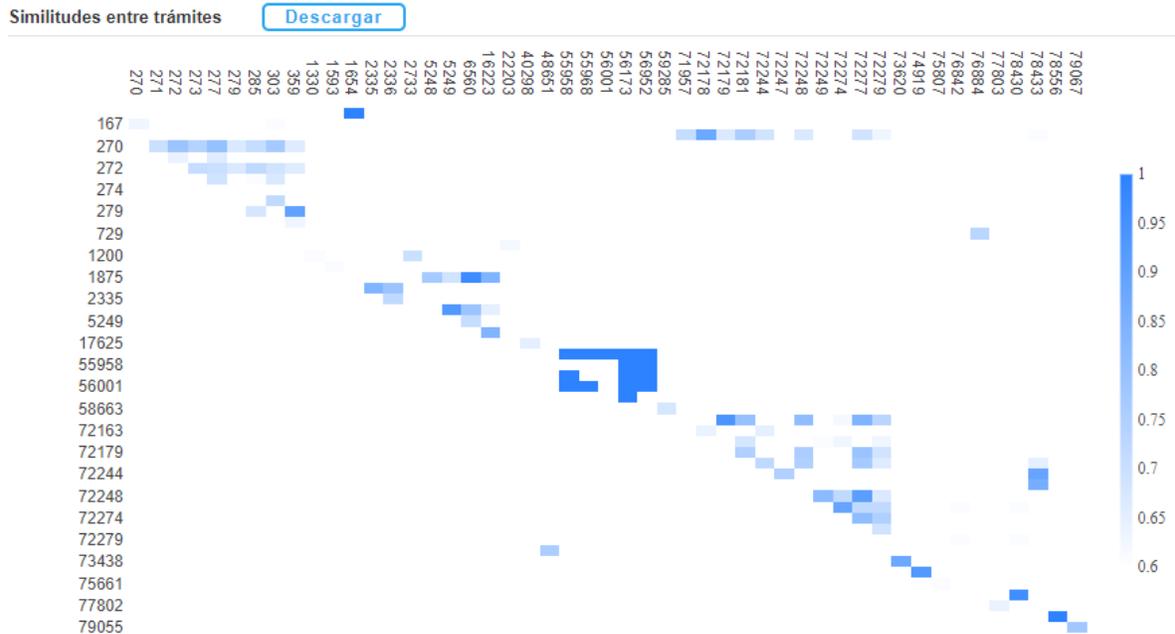
Trámite: 2302	Trámite: 2335
Nombre: Afiliación de embarcaciones	Nombre: Desafiliación de embarcaciones artesanales e industriales
Resultado: Certificado de afiliación de embarcaiones artesanales e industriales	Resultado: Certificado de desafiliación de embarcaciones artesanales e industriales
Entidad: Gobernación de san andres	Entidad: Gobernación de san andres
Municipio: Providencia	Municipio: Providencia
Departamento: Archipiélago de San Andrés, Providencia y Santa Catalina	Departamento: Archipiélago de San Andrés, Providencia y Santa Catalina
Sector: No aplica	Sector: No aplica
Región PND: Caribe	Región PND: Caribe
Totalmente en línea: No Duración: en 15 Día(s) - hábil(es) Requiere pago: No	Totalmente en línea: No Duración: en 15 Día(s) - hábil(es) Requiere pago: No
Documentos: Documentos comunes: <ul style="list-style-type: none"> • Carta de solicitud • Cédula de ciudadanía • Cédula de extranjería Documentos diferentes: <ul style="list-style-type: none"> • Plan de actividades o de trabajo • Registro de matrícula de la embarcación ante capitanía de puerto 	Documentos: Documentos comunes: <ul style="list-style-type: none"> • Carta de solicitud • Cédula de ciudadanía • Cédula de extranjería Documentos diferentes:

Fuente: Elaboración propia.



3. Mapa de similitudes entre trámites empresariales, que permite comprar de forma sencilla los trámites que comparten semejanzas en su descripción mayores al 60%. En la Figura 12 se muestra la visualización del mapa de similitudes, entre más semejante sea la descripción entre dos trámites, se muestra un color azul oscuro (similitud cercana a 1 o 100%) y a medida que esta semejanza disminuye el color se va atenuando (solo se muestran las similitudes mayores al 60%).

Figura 12: Mapa de similitudes entre trámites empresariales dentro de la herramienta de visualización.



Fuente: Elaboración propia.

Nota: Esta herramienta de visualización se puede consultar desde la intranet del DNP en: <http://vdatascience:8045/>

5. Conclusiones y recomendaciones

A partir de la metodología desarrollada y de los resultados obtenidos para cumplir los objetivos de este proyecto, planteados en el plan de trabajo, se presentan a continuación las principales conclusiones obtenidas por el equipo de la UCD y las principales recomendaciones para un mejor uso y aprovechamiento del proyecto.

1. Se evidencia a través de los análisis descriptivos de las bases de datos que son muy pocas las entidades territoriales que han reportado los trámites que van dirigidos para empresas. La herramienta de comparación de trámites puede ser un insumo importante para identificar a estas entidades y a trámites que no se han identificado.
2. Los trámites empresariales que reportan mayor número de PRDS y sensaciones de disgusto (dislikes) son trámites ofrecidos por la Dirección de Impuestos y Aduanas Nacionales, este patrón se mantenido en los dos últimos años.
3. Las técnicas de web-scraping para conseguir información adicional de los trámites ayudó a corregir información incompleta que se reporta en la base de datos “Información registrada en el SUIT de los trámites u otros procedimientos administrativos ([ver](#))” que se sube al portal de datos.gov.co. Se recomienda tener cuidado con los trámites que cuentan descripciones con mucho texto, pues la información se está perdiendo.
4. Se recomienda subir las bases de datos de interacciones (likes/dislikes) a un portal que se pueda consultar de forma automática, al momento, la herramienta funciona con la versión compartida por el equipo de Función Pública con corte a marzo 31 del 2021.



6. Socialización

Este proyecto y sus resultados ha sido socializado con el equipo de la Subdirección de Productividad, Internacionalización y Competencia de la Dirección de Innovación y Desarrollo Empresarial, funcionarios de Función pública, al interior de la Dirección de Desarrollo Digital y la Subdirector General Sectorial del DNP.

Contacto

Si tiene alguna duda, comentario o sugerencia sobre este proyecto, o si le gustaría conversar con la Unidad de Científicos de Datos sobre la posibilidad de una nueva fase para el mismo, puede comunicarse con nosotros a través del correo electrónico ucd@dnpgov.co.