



**El futuro  
es de todos**

**DNP**  
Departamento  
Nacional de Planeación



El futuro  
es de todos

DNP  
Departamento  
Nacional de Planeación

# Rastreo de inversión destinada a la economía del cuidado con enfoque de género

**Unidad de Científicos de Datos**  
Dirección de Desarrollo Digital

Agosto, 2019



# Política pública de cuidado

La construcción de la política pública de cuidado puede enriquecerse con la identificación de los programas y proyectos sobre economía del cuidado con enfoque de género



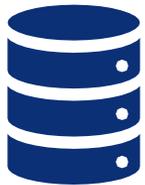
## Objetivo

Caracterizar los programas y proyectos registrados en SUIFP y ACP que pueden estar relacionados con la construcción de la política pública de cuidado.



## Metodología

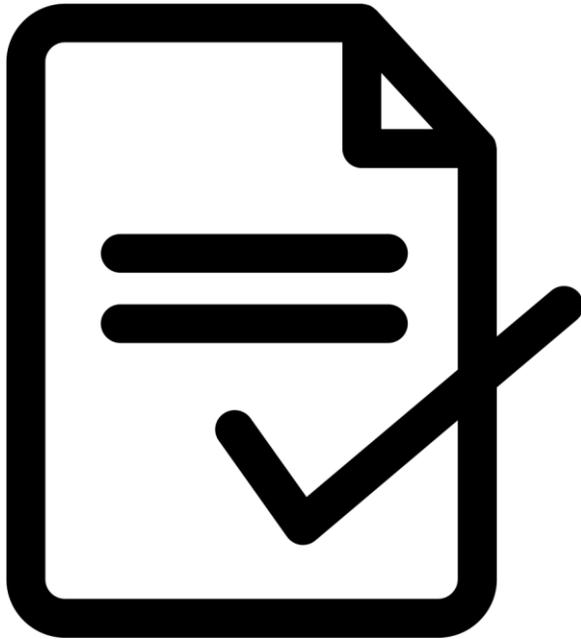
1. Técnicas de procesamiento y análisis de texto.
2. Modelos automáticos para clasificar textos con características similares.



## Insumos

1. 111.156 proyectos de SUIFP con corte a Junio de 2019 (103,899 de regalías y 7257 de PGN).
2. 886 proyectos de APC.
3. 45 documentos de política pública.

# Preprocesamiento de texto



1

Eliminación de números, signos de puntuación y caracteres especiales

2

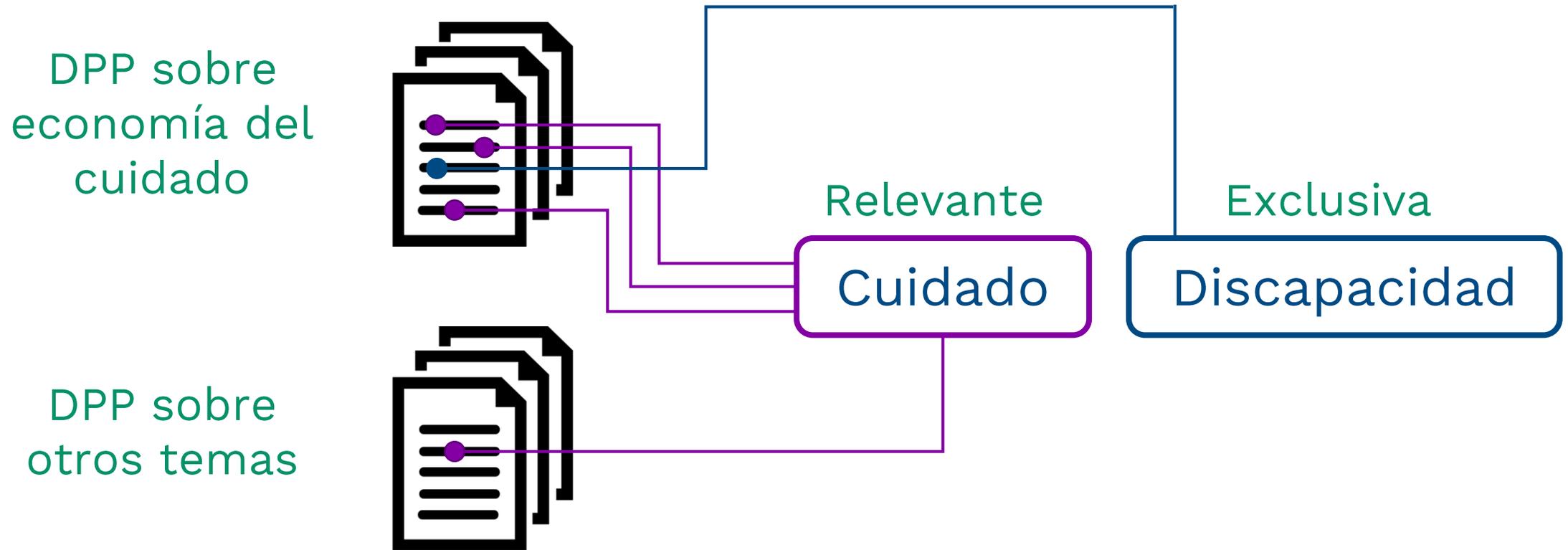
Eliminación de zonas geográficas y palabras irrelevantes (*stopwords*)

3

Conversión de palabras a sus lemas:  
géneros → género  
genero → generar

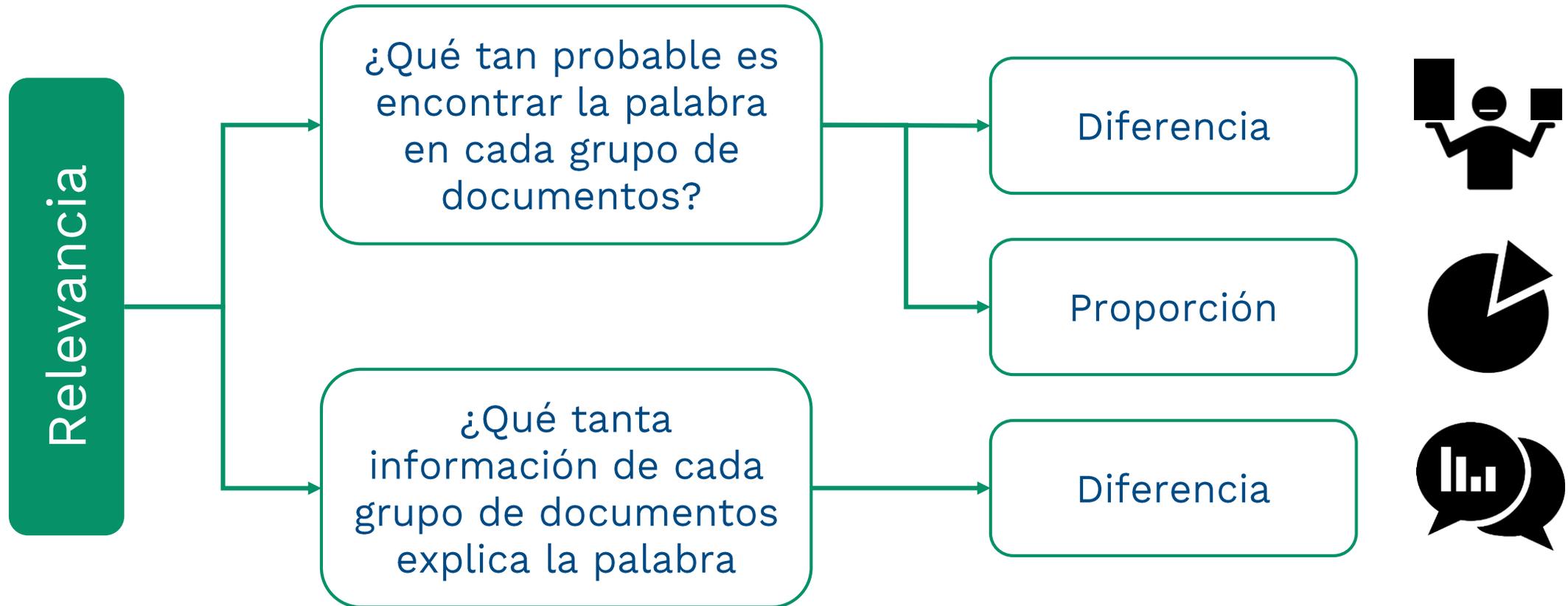
# Selección y comparación de documentos

Se seleccionaron 36 documentos de política pública (DPP) de 12 sectores económicos, los cuales se compararon con 3 DPP sobre equidad de género y 6 sobre cuidado.



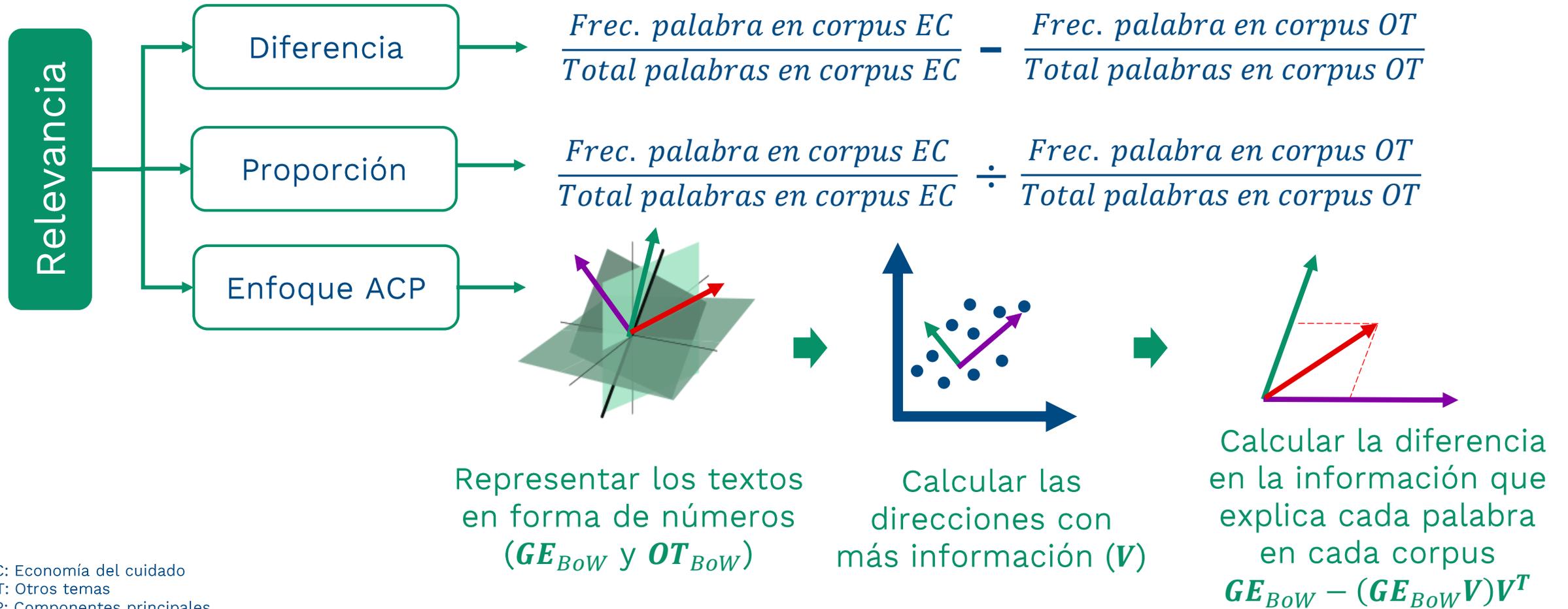
# Identificación automática de palabras clave

Se calcularon tres indicadores de relevancia para cada palabra.



# Identificación de palabras clave

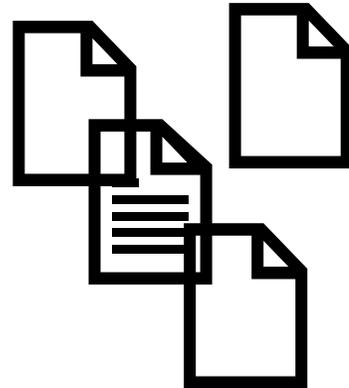
Se calcularon tres indicadores de relevancia para cada palabra.



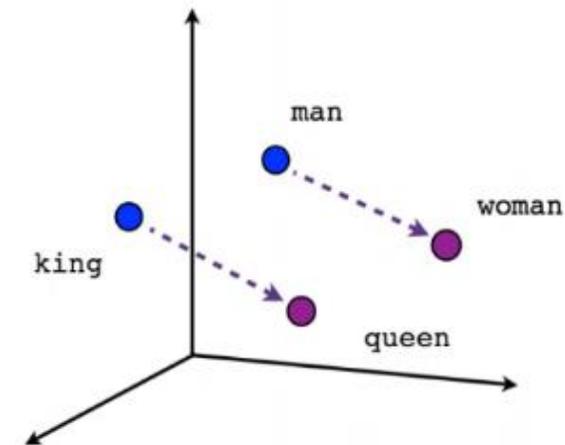
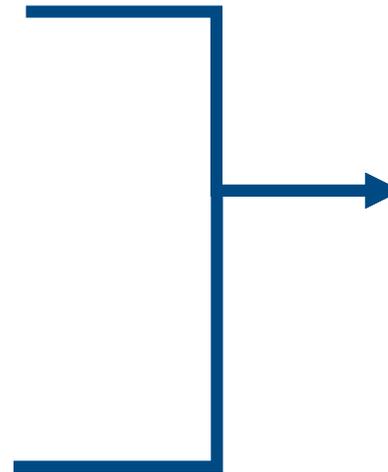
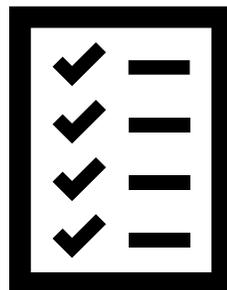
# Inclusión de sinónimos

Como el lenguaje de los proyectos y el de los documentos no es el mismo, se utilizó un algoritmo adicional para identificar sinónimos de las palabras obtenidas.

Corpus de  
Textos



Palabras  
clave



Medición de similitud  
semántica con Word2vec



# Puntuación

Se dio a cada proyecto un puntaje con base en las palabras presentes en su título, descripción, problemática y objetivos.



Metodología para asignar puntajes a los textos de los proyectos

# Visualización

DNP - UCD
☰

**Origen Proyectos**

SUIFP

---

**Incluir proyectos con inicio entre:**

2,004 2,017 – 2,019 2,020

2,004 2,008 2,012 2,016 2,020

---

**Puntaje de proyectos**

Palabras presentes

---

Ver descripción

---

**Filtrar**

No filtrar

---

**Umbral de clasificación (puntaje):**

0 45 171

0 17 34 51 69 86 103 120 137 154 171

---

[Descargar clasificados](#)

Clasificación
Inversión
Módulo

**Ver proyectos clasificados:**

Primeros  Últimos

**Mostrar proyectos**

3

**Proyectos clasificados con mayor puntaje**

N.	Origen	Bpin	Nombre_proyecto	Sector
1	1004000440000	SUIFP PGN	ASISTENCIA A LA NINEZ Y APOYO A LA FAMILIA PARA POSIBILITAR EL EJERCICIO DE LOS DERECHOS	INCLUSIÓN SOCIAL Y RECONCILIACIÓN
2	0050002340000	SUIFP PGN	IMPLANTACION PROGRAMA DE APOYO INTEGRAL A MUJERES JEFAS DE HOGAR	PRESIDENCIA
3	1004001060000	SUIFP PGN	PROTECCION -ACCIONES PARA PRESERVAR Y RESTITUIR EL EJERCICIO INTEGRAL DE LOS DERECHOS DE LA NINEZ Y LA FAMILIA -ATENCION A LA POBLACION DESPLAZADA -APD A NIVEL NACIONAL	INCLUSIÓN SOCIAL Y RECONCILIACIÓN



Esta técnica permitió identificar proyectos de inversión relacionados con la economía del cuidado.



El uso del algoritmo, complementado con el conocimiento experto, permitió asignar puntajes a los proyectos para determinar su nivel de relación con el tema de economía del cuidado.



Con la interfaz, es posible redefinir fácilmente algunos criterios de clasificación, en caso de considerarse pertinente.



## Conclusiones



**El futuro  
es de todos**

**DNP**  
Departamento  
Nacional de Planeación