



**El futuro
es de todos**

DNP
Departamento
Nacional de Planeación



El futuro
es de todos

DNP
Departamento
Nacional de Planeación




Análisis de palabras clave en documentos CONPES por período de gobierno

Unidad de Científicos de Datos
Dirección de Desarrollo Digital

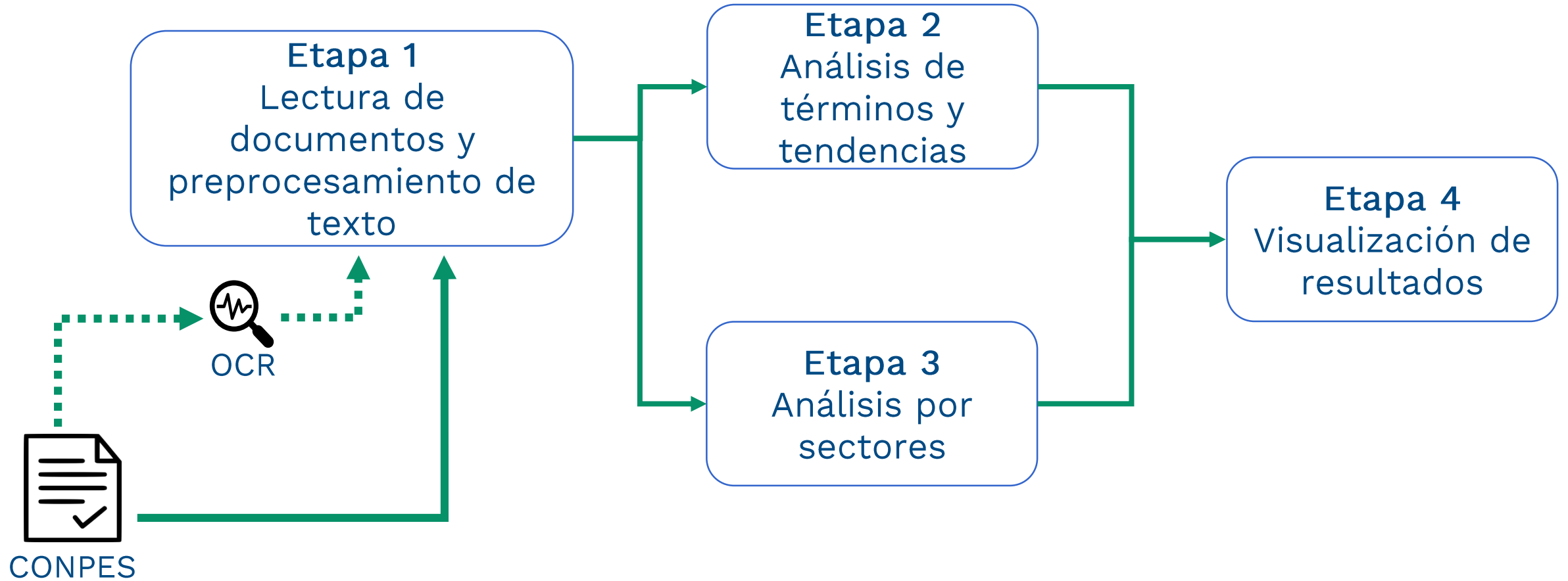
Octubre, 2020



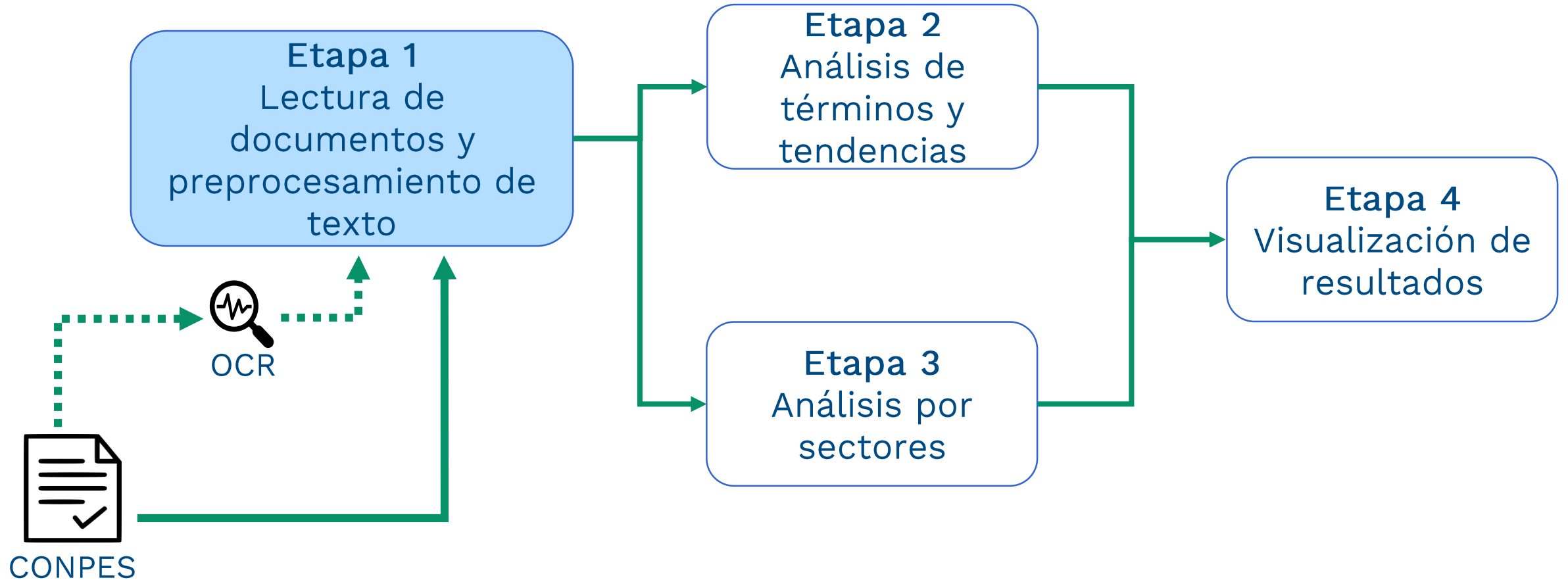
Análisis de los CONPES a través del tiempo

¿Qué?	Identificar los temas relacionados a la política pública del país que tienen mayor participación y relevancia durante los distintos períodos de gobierno a través de los documentos CONPES	
¿Para qué?	Para conocer la evolución en el tiempo de los lineamientos de política pública y orientar la construcción de los Planes Nacionales de Desarrollo (PND) y de las políticas públicas	
Resultado	Aplicativo de visualización que contiene: <ol style="list-style-type: none">1. Búsqueda de documentos2. Resultados gráficos interactivos	

Metodología

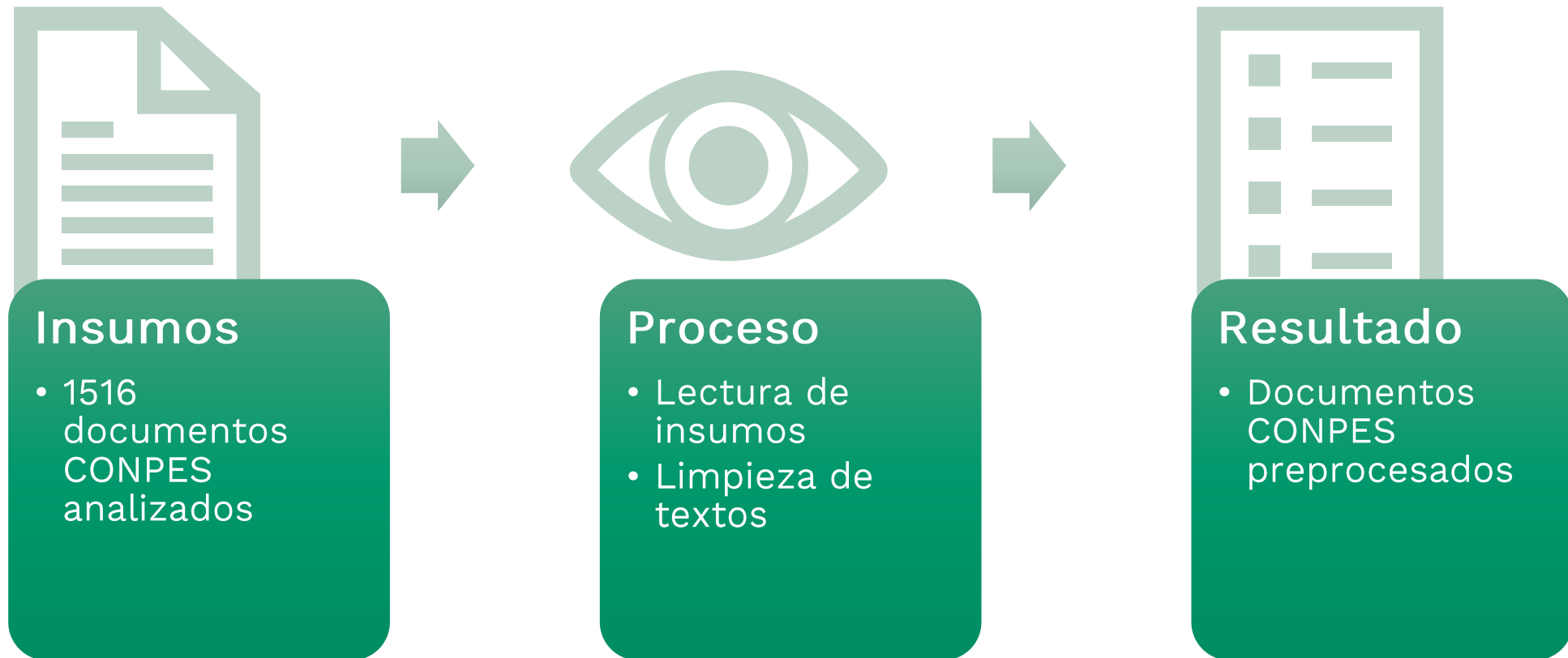


Metodología



Metodología-Etapa 1

Lectura de documentos y preprocesamiento de texto.



Metodología

Etapa 1: lectura de documentos y preprocesamiento de texto.

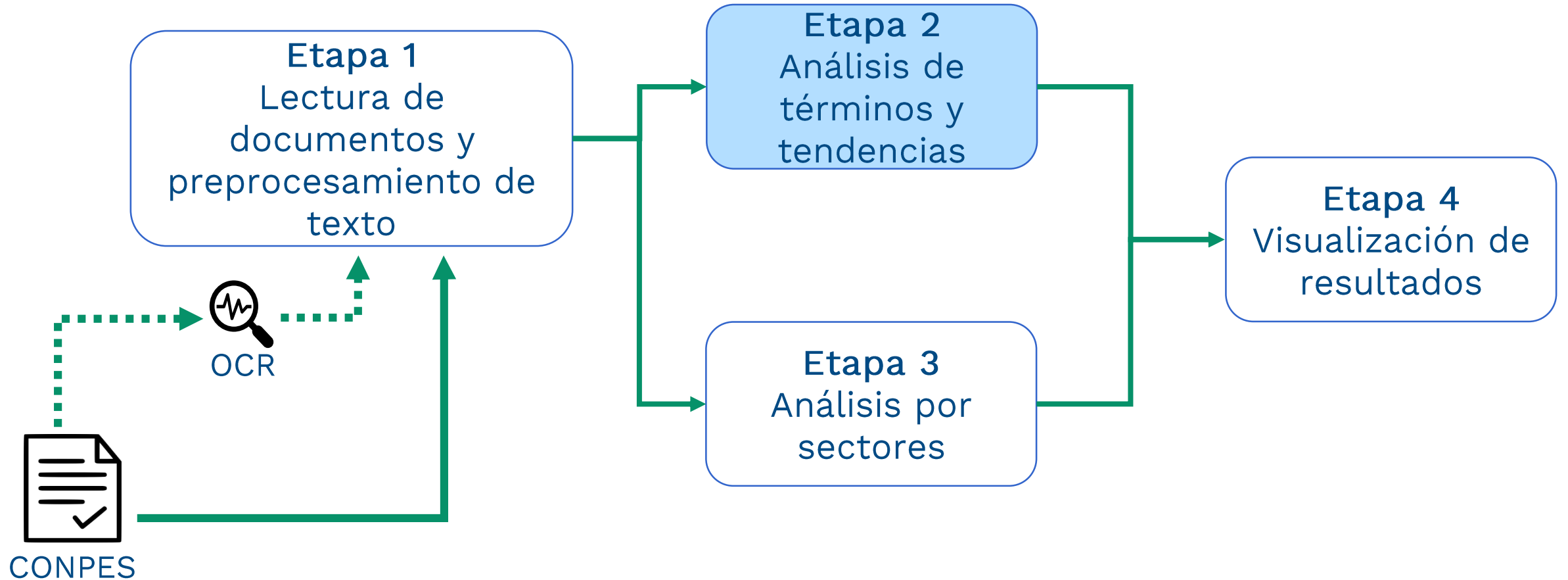


Grupo	Cantidad
Documentos sin link (excluidos)	49
Documentos leídos sin OCR	1266
Documentos leídos con OCR	201
Total	1516



- Caracteres especiales
- Números
- Signos de puntuación
- Stopwords*

Metodología



Metodología-Etapa 2

Análisis de términos y tendencias



Insumos

- Documentos CONPES procesados



Proceso

- Identificación de palabras clave
- Análisis de coocurrencias




Producto

- Nubes de palabras y bigramas
- Red de términos
- Por periodo presidencial

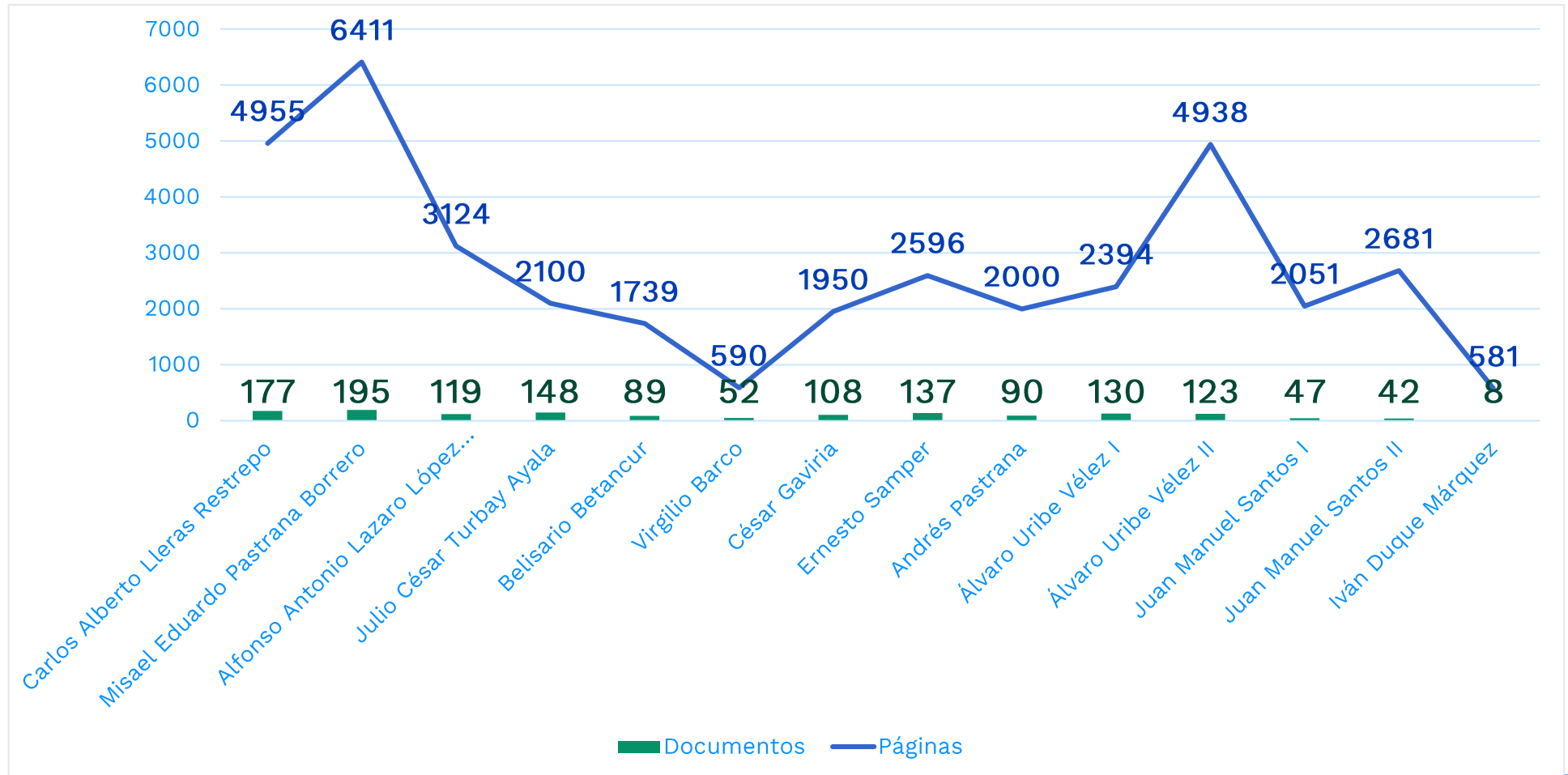
Metodología-Etapa 2

Análisis de términos y tendencias



Insumos

- Documentos CONPES procesados



Carlos Alberto Lleras Restrepo



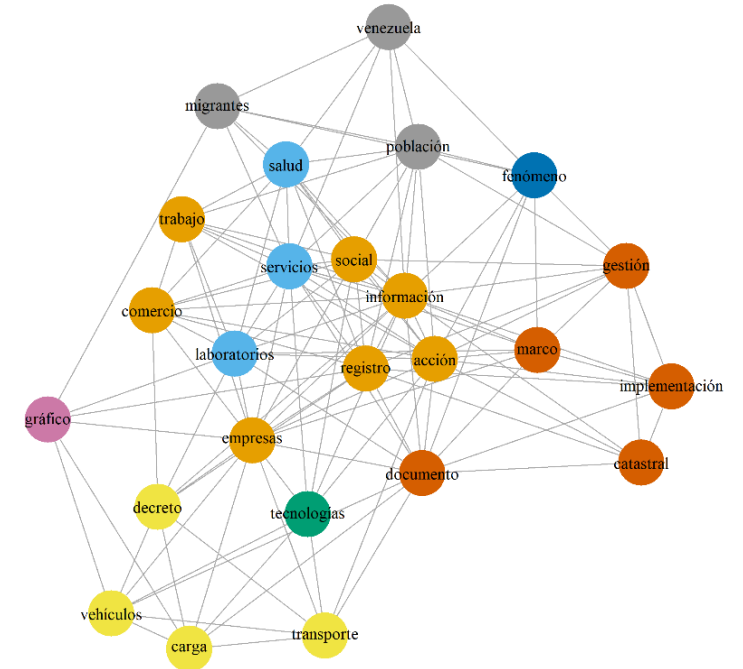
Nube de palabras

Andrés Pastrana



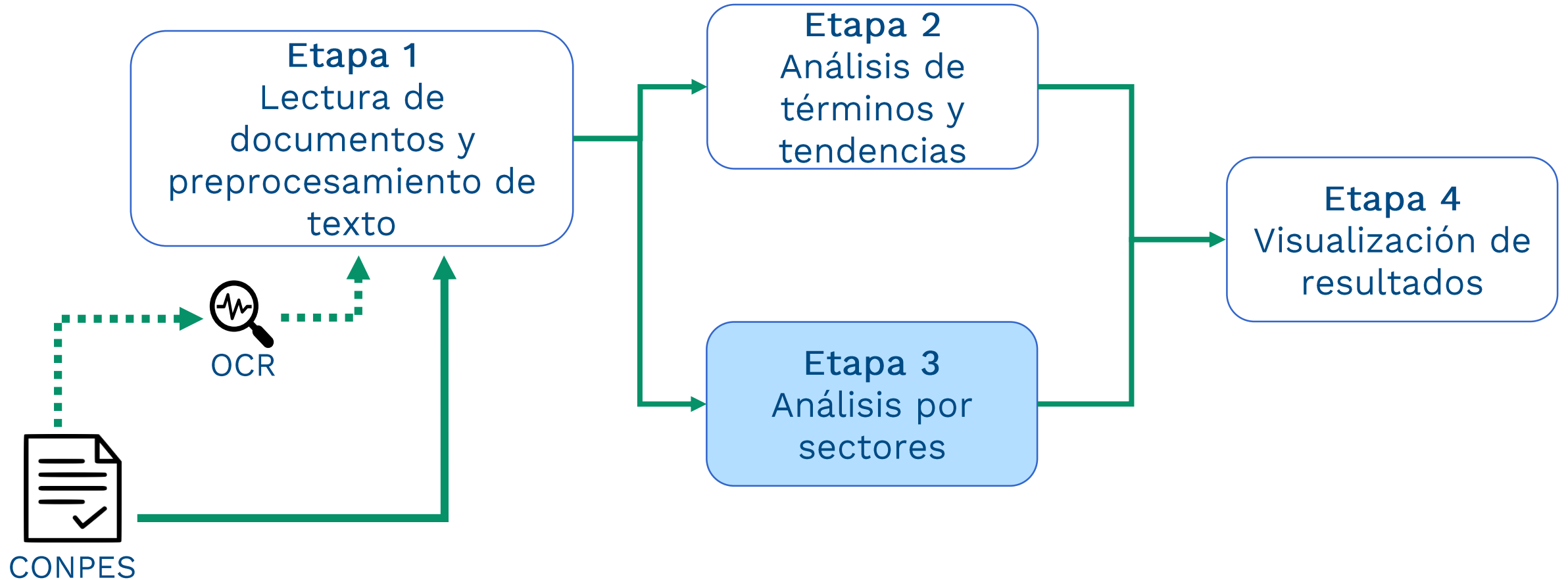
Nube de bigramas

Iván Duque Márquez



Co-ocurrencias

Metodología



Metodología-Etapa 3

Análisis por sectores

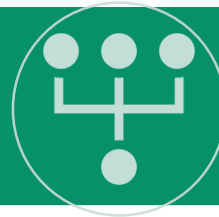
- CONPES etiquetados por sector

Insumos



- Bag of words
- Vectorización
- Palabras ambiguas

Proceso



- Tendencias de términos por sectores a través del tiempo

Producto

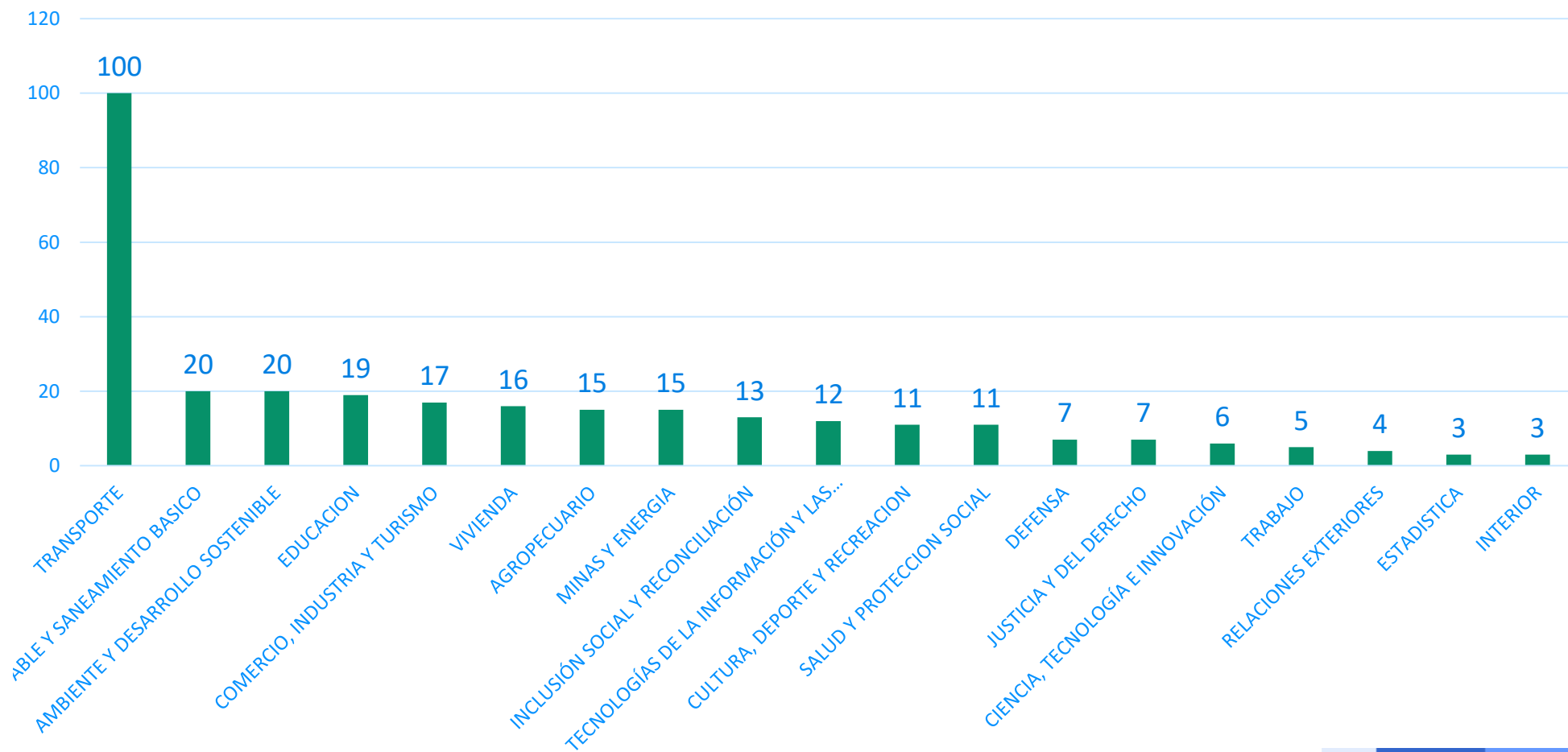


Metodología-Etapa 3

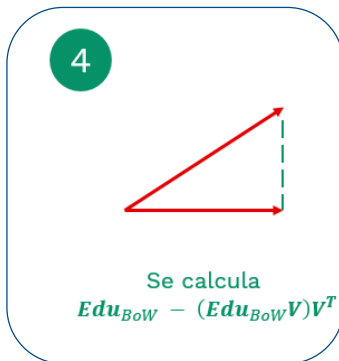
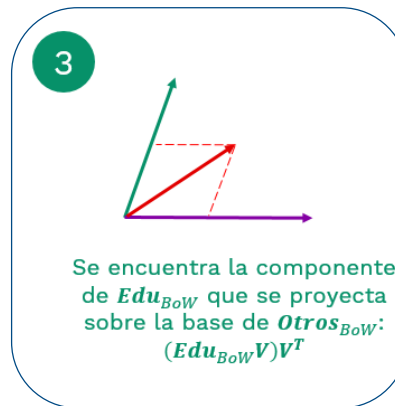
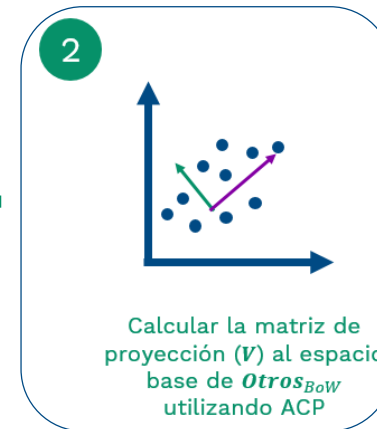
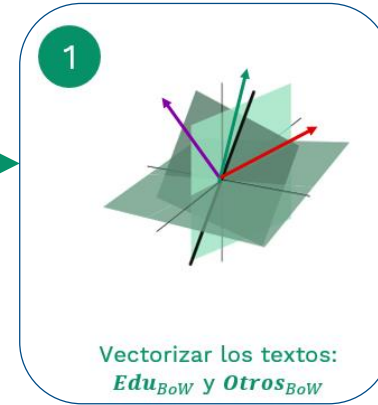
Análisis por sectores

- CONPES etiquetados por sector

Insumos 



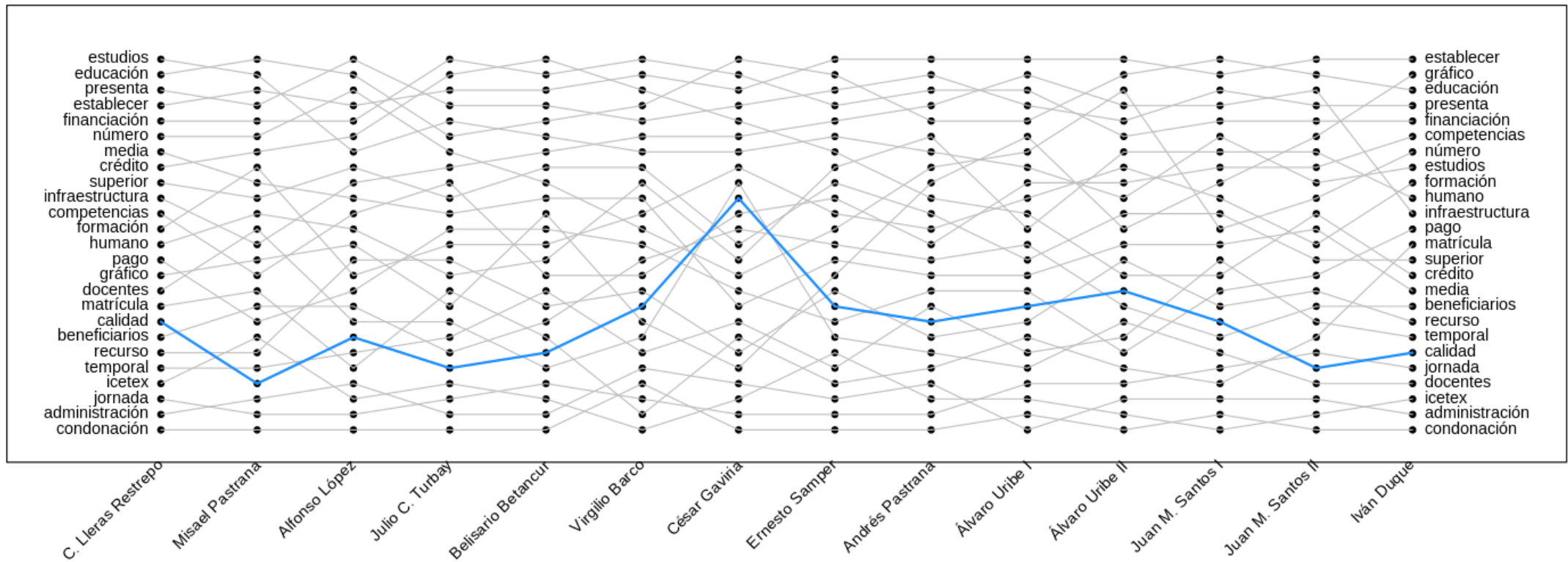
Bag of words	un	texto	ejemplo	otro	de	más
un texto ejemplo	1	1	1	0	0	0
otro texto ejemplo	0	1	1	1	0	0
ejemplo de texto	0	1	1	0	1	0
un texto más texto	1	2	0	0	0	1

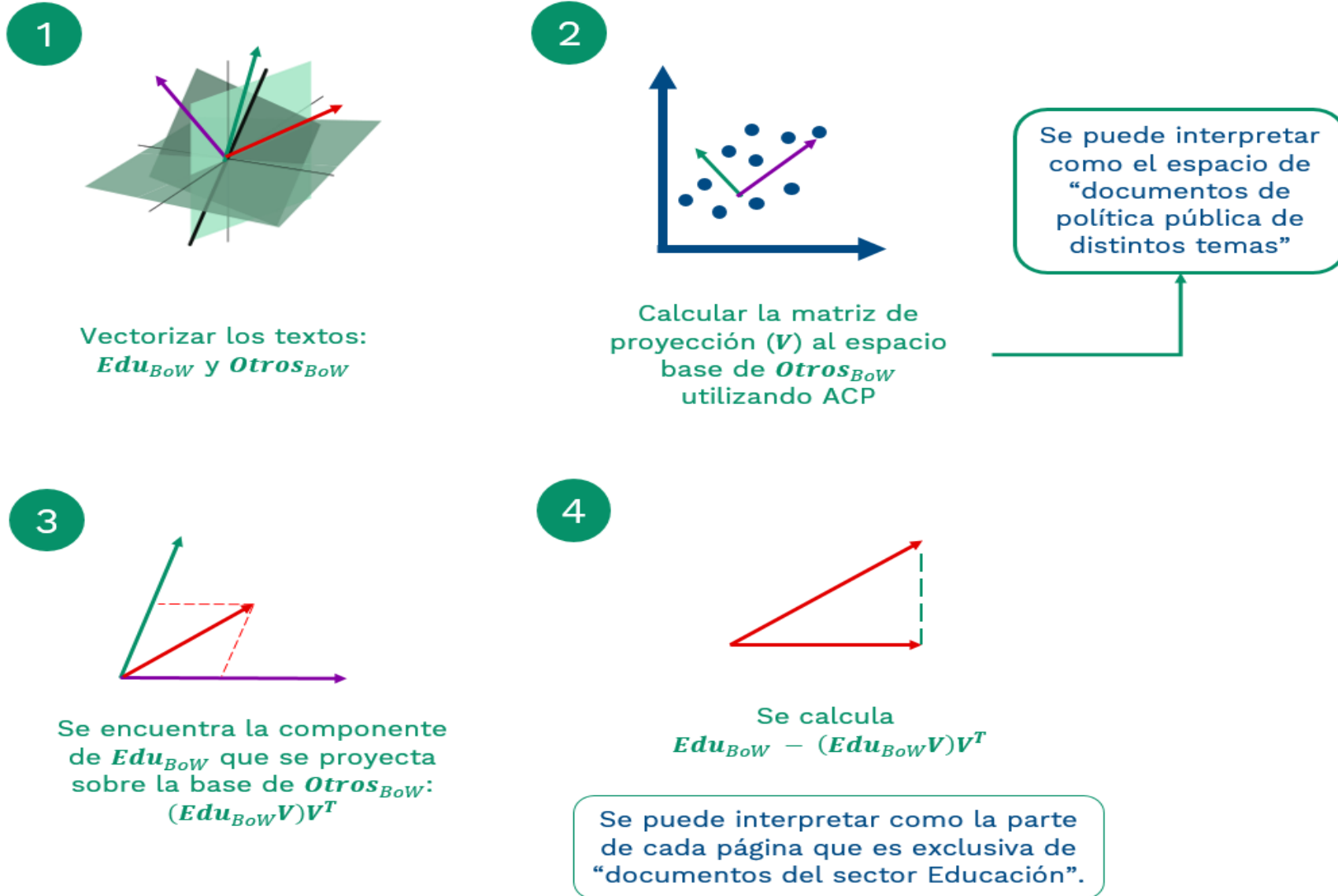


Metodología-Etapa 3

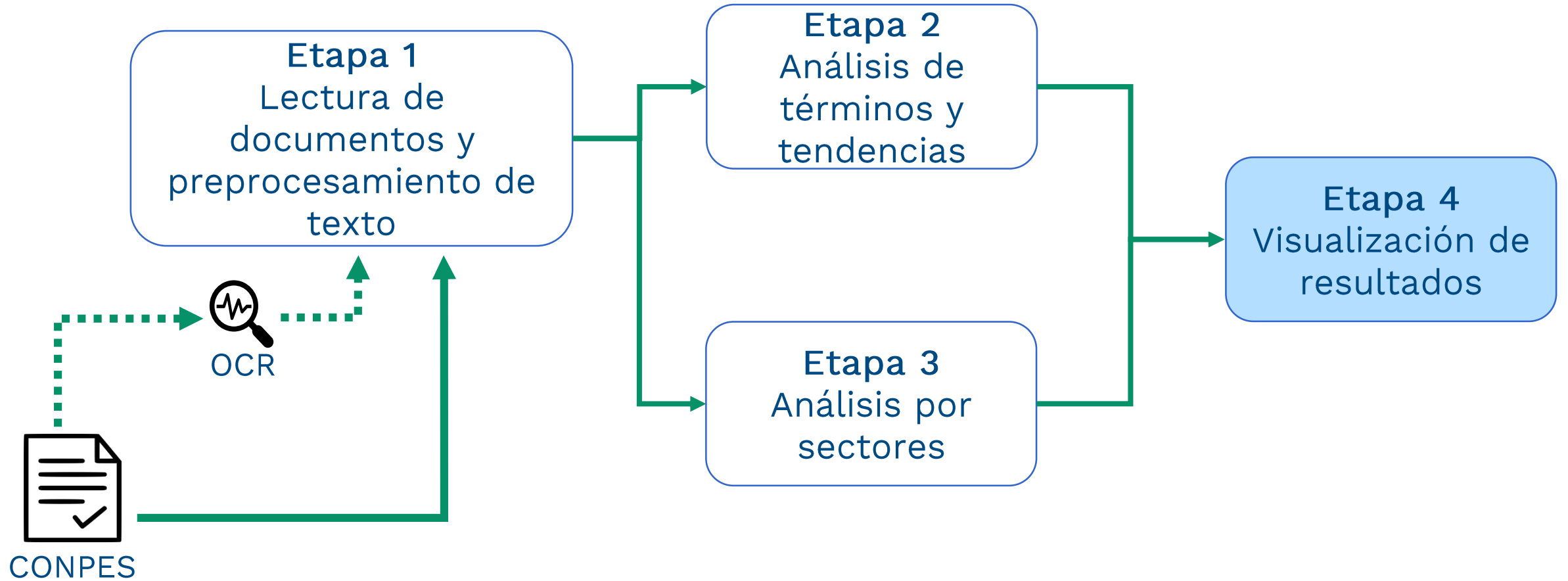
Análisis por sectores

EDUCACION





Metodología



Metodología – Etapa 4

Visualización de resultados



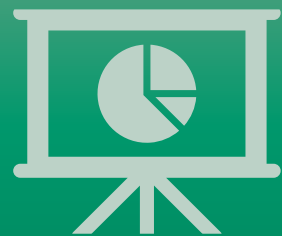
Insumos

- Documentos CONPES preprocesados
- Gráficos descriptivos y de tendencias



Proceso

- Diseño de la aplicación con el paquete Shiny de R



Resultados

- Herramienta de visualización interactiva

Aplicativo

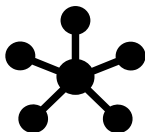
http://10.10.21.20:3838/analisis_conpes/



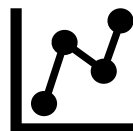
Buscador de documentos



Nubes de palabras y bigramas



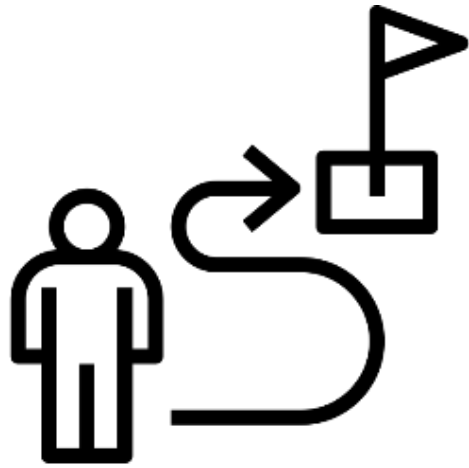
Redes de términos



Tendencias por sector



Conclusiones

**1**

A través de los términos relevantes identificados por periodo presidencial, se puede evidenciar la diferencia del enfoque que dan los gobiernos a sus políticas

2

Se puede realizar un seguimiento del cumplimiento que dan los gobiernos a la solución de problemáticas y a lo establecido en sus planes de gobierno

3

Se observa la importancia que da cada gobierno a los diferentes sectores económicos

Con los resultados de este proyecto se facilitará la labor de búsqueda de documentos CONPES por temas específicos, la identificación de los temas relevantes relacionados con la política pública durante los diferentes periodos de gobierno y de esta forma se podrá orientar la construcción futura de los PND y en general de las políticas públicas del país.





**El futuro
es de todos**

DNP
Departamento
Nacional de Planeación