**Informe SICSS 2025: Análisis Sociolingüístico del Plan de Desarrollo de Antioquia (2020–2023)**

**Introducción**

Este informe sintetiza el trabajo realizado en el marco del curso SICSS Bogotá 2025, combinando herramientas de procesamiento de lenguaje natural (PLN) en Python con un análisis sociológico inspirado en la crítica al desarrollo de Arturo Escobar. El estudio se centra en el Plan de Desarrollo Departamental de Antioquia 2020–2023 y utiliza técnicas computacionales para identificar patrones discursivos, desigualdades simbólicas y estructuras de representación social. Se complementa con los informes oficiales de la Gobernación y los programas de inversión, especialmente en juventud e infancia.

**Corpus y enfoque metodológico**

Se utilizó el archivo completo del Plan de Desarrollo como corpus base, al que se aplicó el siguiente pipeline de análisis en Python (código disponible en el archivo `sicss\_pd\_antioquia\_(1) (2).ipynb`:

1. Carga del texto y limpieza: eliminación de puntuación, dígitos, stopwords y lematización con `spaCy`.

2. Análisis temático: agrupación de palabras por categorías sociológicas: clase social, género, etnia, grupo etario, institucionalidad y territorio.

3. Visualización: gráficas de barras por eje temático.

4. Nube de palabras: representación del campo semántico del texto.

5. Análisis de sentimientos: usando modelos de Transformers multilingües (`nlptown/bert-base-multilingual-uncased-sentiment`). Este proceso permitió construir una lectura discursiva del desarrollo departamental desde una visión sociocrítica.

**Hallazgos clave**

1. Tópicos y frecuencias temáticas

Institucionalidad: palabras como “programa”, “gestión”, “modelo”, “plan” dominan el discurso, lo cual indica una centralidad del aparato técnico-estatal (Foucault, 2007).

Territorio: aparecen “municipio”, “rural”, “subregión” con alta frecuencia, reflejando el enfoque descentralizado del gobierno.

Género y etnia: muy baja presencia de términos como “mujeres”, “afrodescendiente” o “indígena”, lo que evidencia omisiones discursivas significativas (Fraser, 2006).

**2. Nube de palabras**

Se elaboró una nube de palabras que incluyó todo el texto lematizado sin stopwords. Las más relevantes fueron: `desarrollo`, `territorio`, `equidad`, `gobierno`, `niños`, `programa`, `gestión`, `vida`, `pobreza`, `educación`. Esto sugiere un discurso centrado en bienestar técnico pero con vacíos simbólicos hacia la diversidad (Bourdieu, 1997) (Escobar, 2007)

**3. Análisis de sentimientos**

La mayoría de los fragmentos del plan se calificaron como \*\*positivos o neutros\*\*. Este predominio refleja una narrativa institucional optimista que busca legitimidad pública. Términos emocionales como “vida”, “progreso” y “unidos” tienen función retórica y persuasiva, construyendo una visión hegemónica del desarrollo (Foucault, 2007).

**4. Inclusión y desigualdad desde los datos de inversión**

A partir del cuadro comparativo y los informes oficiales (`inversión departamental `, `informe-gobernacion-baja-dic11`), se identificaron patrones de distribución presupuestal:

Alta inversión en “actores clave” (más de 26.000 personas) y programas con enfoque en infancia y adolescencia.

Bajo detalle sobre población étnica o comunidades rurales en el desglose de beneficiarios.

Visión cuantitativa centrada en cobertura y ejecución, sin claridad sobre impactos diferenciales.

Este hallazgo refuerza lo observado en el análisis léxico: el discurso del desarrollo reproduce lógicas de intervención sin representar completamente las subjetividades de los grupos excluidos.

**Reflexión social**.

Desde la crítica posdesarrollista de Escobar (2007), el lenguaje técnico del Plan proyecta una imagen única y modernizante del desarrollo. Este se manifiesta como una “tecnología de gobierno” (Foucault, 2007), que excluye epistemologías diversas (indígenas, populares, rurales) en favor de una racionalidad planificadora.

La evidencia empírica extraída mediante PLN corrobora esta hipótesis: predominan los dispositivos lingüísticos de planificación, eficiencia, gestión y orden territorial, mientras que las diferencias estructurales (género, etnia, clase) se diluyen o se mencionan marginalmente. Desde Bourdieu (1997), esto se traduce en una reproducción de capital simbólico institucional que refuerza jerarquías sociales. Fraser (2006), por su parte, subraya cómo la justicia redistributiva sin reconocimiento cultural limita la equidad real del discurso político

Ventajas y limitaciones metodológicas

**Ventajas:**

Python permite escalar el análisis a documentos extensos de manera reproducible.

Las bibliotecas `spaCy`, `transformers`, `matplotlib` y `WordCloud` permiten combinar rigor técnico y claridad visual.

**Limitaciones**:

El análisis semántico superficial no capta ironías, metáforas o tensiones discursivas implícitas.

Modelos de sentimiento no siempre capturan el matiz sociopolítico del texto.

Validación cruzada: se contrastaron hallazgos computacionales con los datos del informe de la Gobernación y el documento de inversión departamental, garantizando una triangulación metodológica entre discurso, presupuesto y acción institucional.

**Conclusión**

Este informe demuestra el valor de integrar el análisis de lenguaje natural en Python con enfoques sociológicos para explorar discursos de poder y desarrollo. El caso de Antioquia revela cómo el lenguaje del plan refuerza ciertas narrativas institucionales mientras omite otras realidades sociales.

A través de los gráficos, nubes y categorización temática, se visibilizan las jerarquías simbólicas del documento. Esta metodología puede escalarse a otros territorios y planes, fomentando una sociología computacional crítica y latinoamericana.

**Referencias**

Bourdieu, P. (1997). Razones prácticas. Sobre la teoría de la acción. Anagrama.

Escobar, A. (2007). La invención del Tercer Mundo. Construcción y deconstrucción del desarrollo. Norma.

Foucault, M. (2007). Seguridad, territorio, población. Fondo de Cultura Económica.

Fraser, N. (2006). ¿De la redistribución al reconocimiento? Dilemas de la justicia en la era postsocialista. Morata.

Tufekci, Z. (2014). Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. In ICWSM.

Gobernación de Antioquia (2023). Informe de Gestión: Plan de Desarrollo Unidos por la Vida 2020–2023.

Gobernación de Antioquia (2023). Inversión Departamental. Secretaría de Planeación.