

## ¿QUE ES UN ANALIZADOR LEXICO?

Un analizador léxico o analizador lexicográfico (en inglés scanner o tokenizer) es la primera fase de un compilador, consistente en un programa que recibe como entrada el código fuente de otro programa (secuencia de caracteres) y produce una salida compuesta de tokens (componentes léxicos) o símbolos. Estos tokens sirven para una posterior etapa del proceso de traducción, siendo la entrada para el analizador sintáctico.

Un token léxico o simplemente token es una cadena con un significado asignado y, por lo tanto, identificado. Está estructurado como un par que consta de un nombre de token opcional. El nombre del token es una categoría de unidad léxica. Los nombres de token comunes son:

**identificador:** nombres que elige el programador.

**palabra clave:** nombres que ya están en el lenguaje de programación.

**separador (también conocidos como signos de puntuación):** caracteres de puntuación y delimitadores emparejados.

**operador:** símbolos que operan sobre argumentos y producen resultados.

**literal:** literales numéricos, lógicos, textuales, de referencia.

**comentario:** línea, bloque (Depende del compilador si el compilador implementa comentarios como tokens, de lo contrario, se eliminará).

Nombre simbólico	Ejemplos de valores de token
identificador	x, color, ARRIBA
palabra clave	si, mientras, retorno
separador	}, (, ;
operador	+, <, =
literal	verdadero, 6,02e23, "música"
comentario	<i>/* Recupera datos de usuario */, // debe ser negativo</i>