

La tokenización divide el texto en unidades significativas llamadas tokens.