Dataset Analysis Report

Dataset Path: outputs\hf_ner_dataset

Dataset Overview:

• Total samples: 7
• Total entities: 76
• Number of splits: 2
• Unique labels: 3

Split Details:

• train: 6 samples
• validation: 1 samples

Label Information:

• O: 62 occurrences
• birth_date_place: 6 occurrences
• name: 8 occurrences

Split Statistics
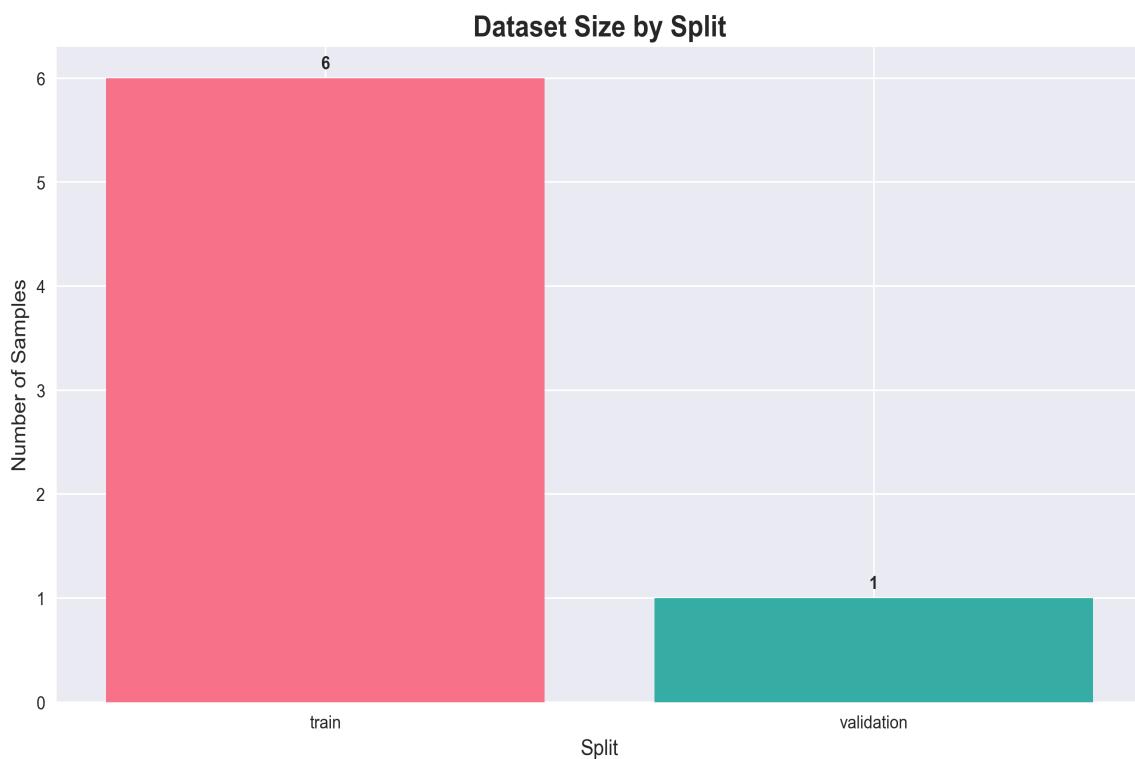
Split	Samples	Entities
train	6	68
validation	1	8

Label Distribution by Split
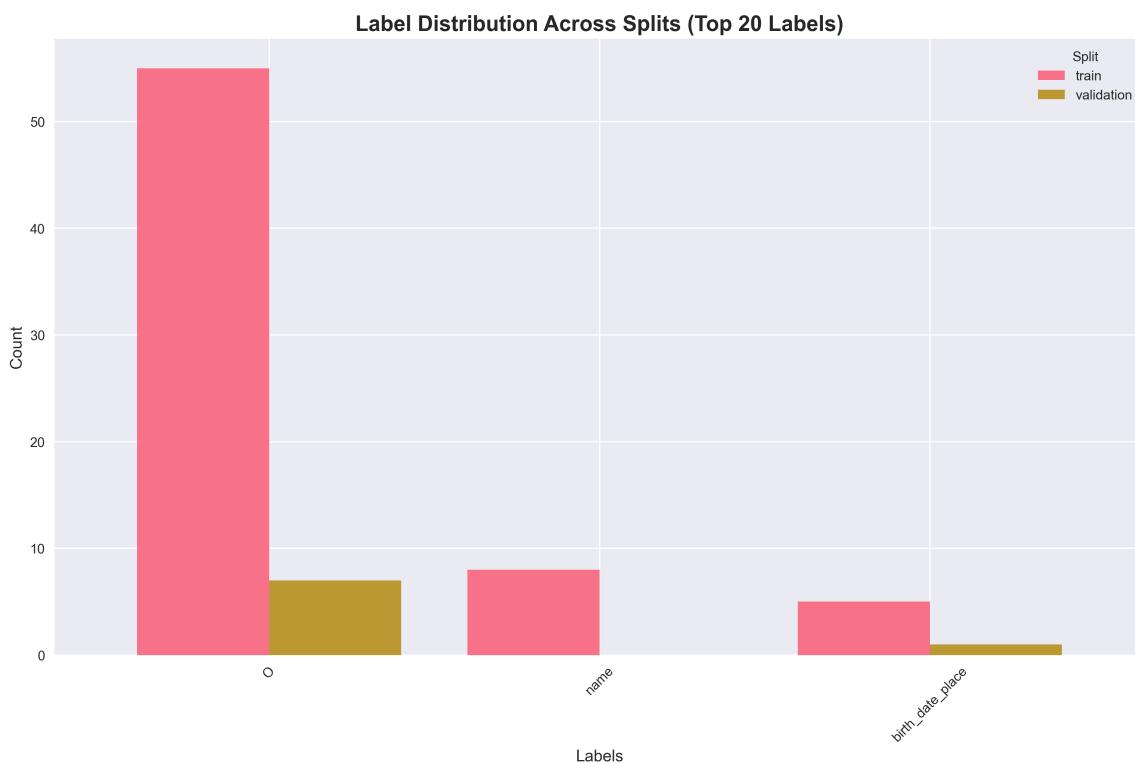
| Label	train	validation
O	55	7
birth_date_place	5	1
name	8	0

Visualizations

Dataset Size Comparison by Split
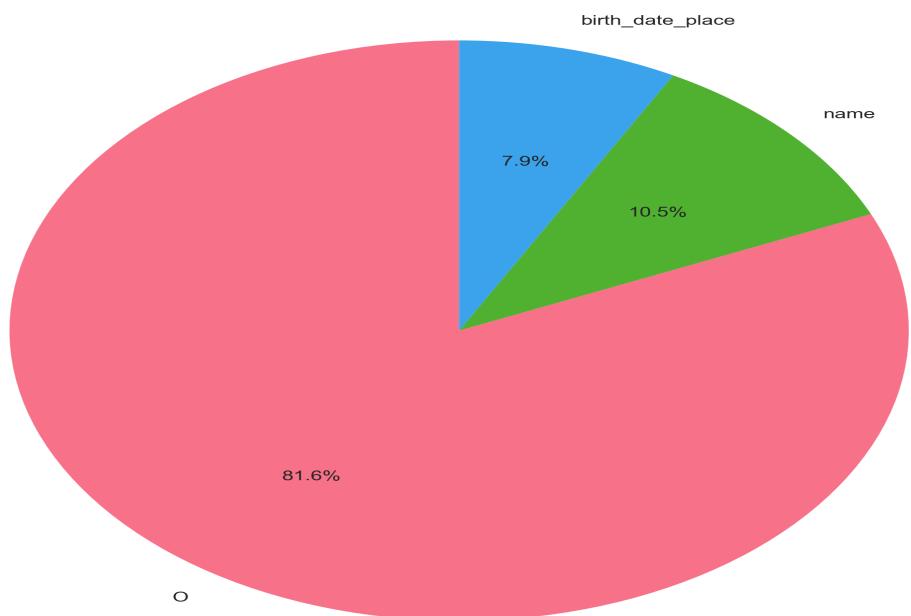
Label Distribution Across Splits (Top 20 Labels)

Overall Label Distribution (Top 20 Labels)

Total Label Frequencies (Top 20 Labels)