

Label

Prediction

Probing model

Span representations

Self-attention

Projection

Concatenated contextual embeddings

Contextual embeddings

Input tokens

