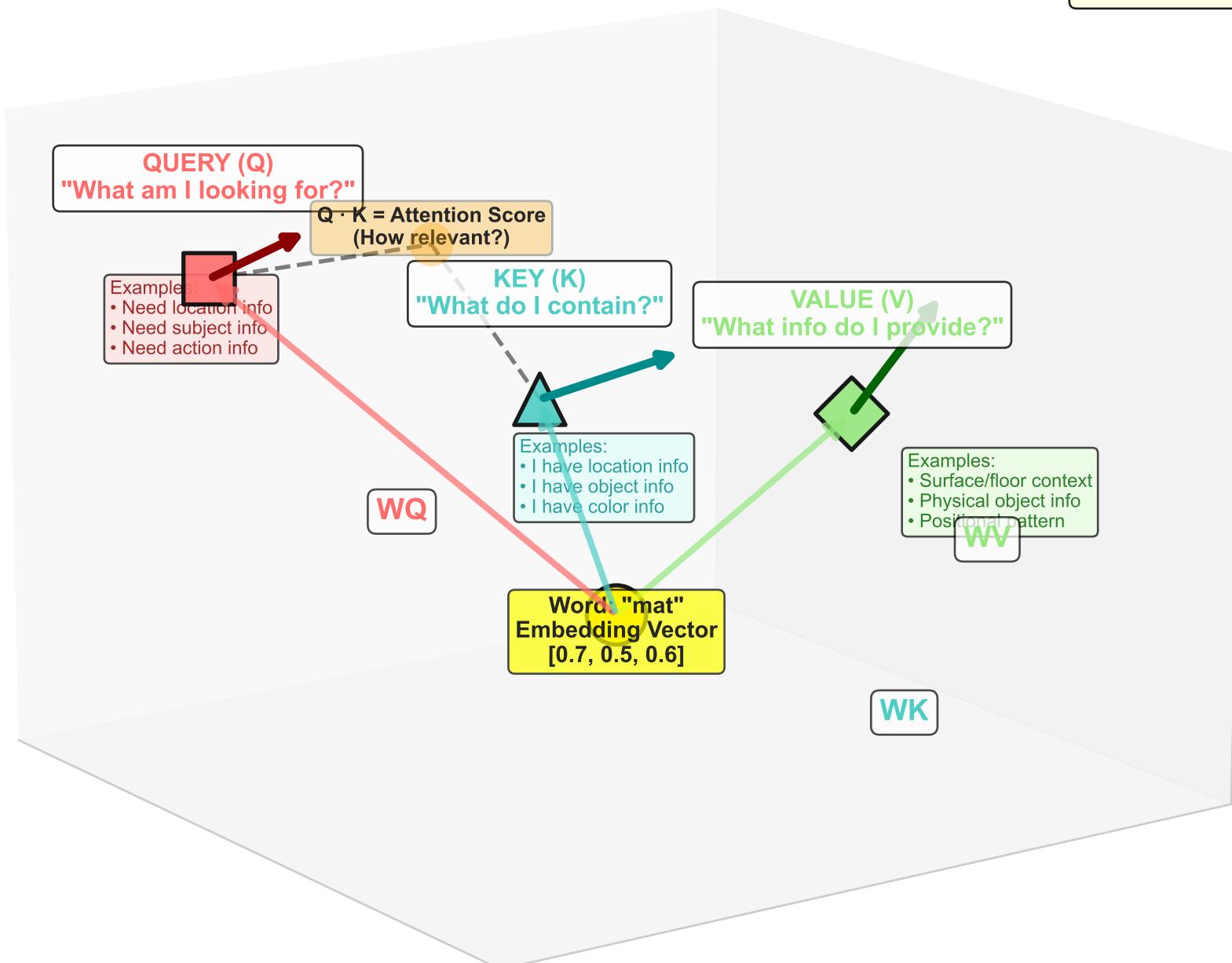


# Query-Key-Value: Three Different Perspectives on Same Word

Each transformation extracts different aspects of meaning

- █ Query: Seeking information
- █ Key: Advertising content
- █ Value: Actual information

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T)V$$



The same word "mat" is transformed into 3 spaces: Q asks what it needs, K advertises what it has, V provides the actual content