## [MASK] is the first transformer-based PLM. **End-to-End Backprop.** - Publimed DDD. **Knowledge Attention |≡**1 Ι≡Ί PubMed Knowledge **Encoder** S2ORC **MIPS FFN Layer Self-Attention Layer** Keys **Values** Wikipedia