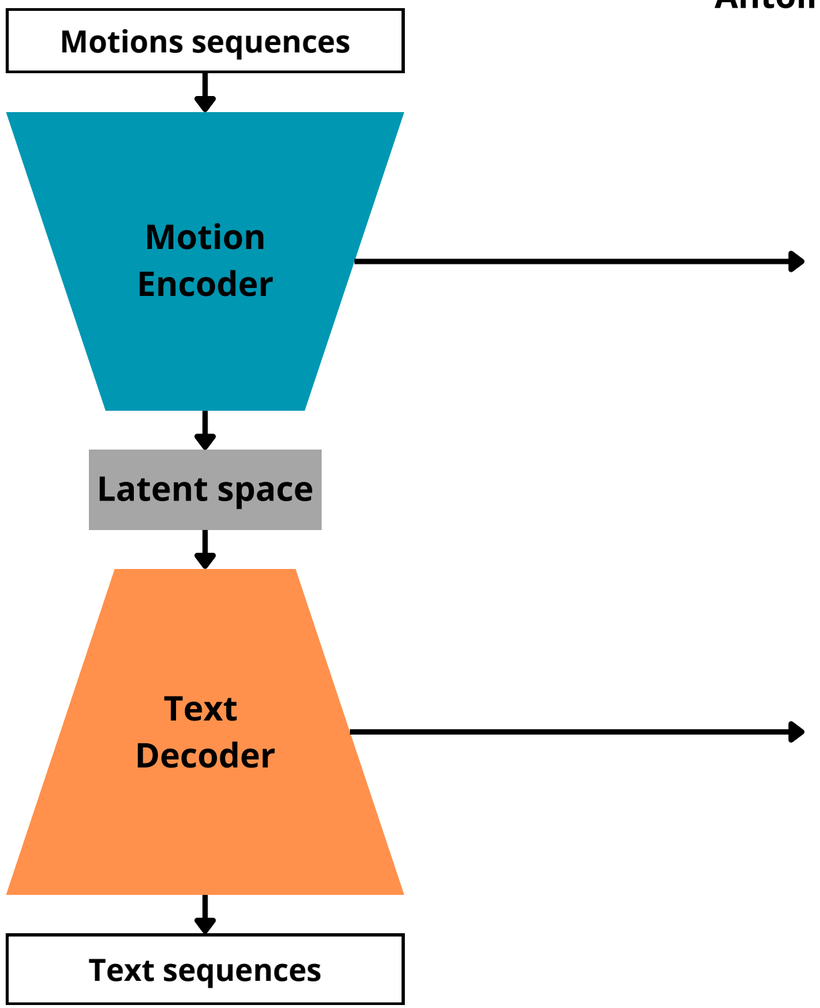
AE M2T

(Auto Encoder Motion to Text)

Antoine Mariot, Léo Delecroix

IMT Nord Europe
École Mines-Télécom
IMT-Université de Lille



- *Features extraction with 1D Conv*
- *Transformer layer for temporal dimension*
- *Fully-Connected layer for GPT2 pipeline*
- *Noise for robustness*

- *Linear projection to adapt GPT2 embeddings*
- *Cross-attention layer to link text and latent vector*
- *Text decoding with GPT2*
- *Parameters optimization (temperature, beams etc...)*

Training :

- *MotionText dataset*
- *Train loader*
- *Cross-entropy loss*
- *Scheduler*
- *GPT2 freeze*
- *Gradient cliping for stabilization*
- *Learning rate evolution*