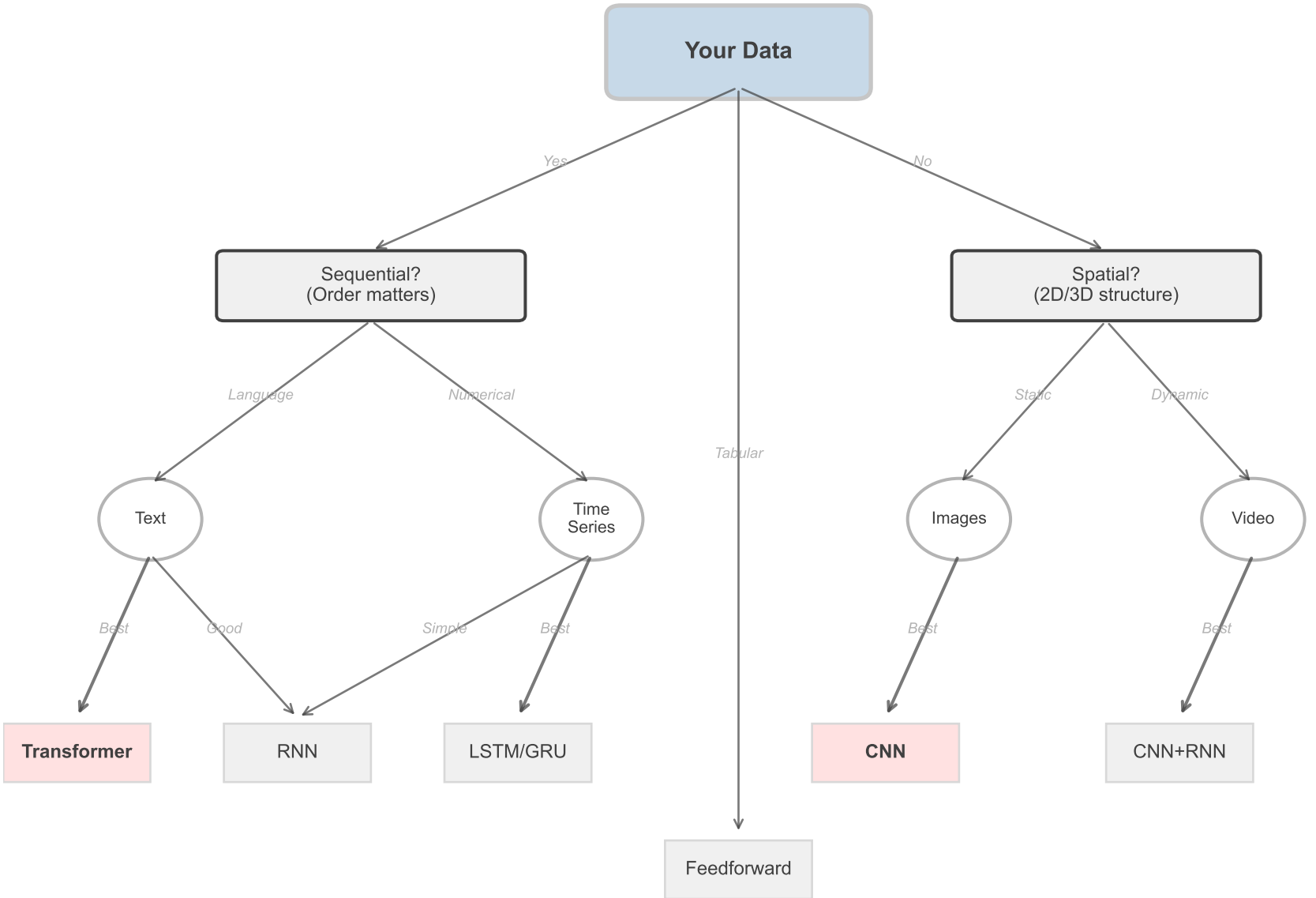
Architecture Selection Guide



** Transformer is now often best for all sequential data*