

# The Integration Of Domain-Specific Code Documentation Data From Stack Overflow Impact The Bleu Score And Functional

Assignee Research

June 7, 2026

## Abstract

This report synthesises findings from 8 peer-reviewed papers addressing the following research question: How does the integration of domain-specific code documentation data from Stack Overflow impact the BLEU score and functional correctness of transformer-based Java code generation models on the. 15 claims were extracted from source literature; 2 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: JaCoText: A Pretrained Model for Java Code-Text Generation. Research question: How does the integration of domain-specific code documentation data from Stack Overflow impact the BLEU score and functional correctness of transformer-based Java code generation models on the CONCODE dataset?.

## 2 Methodology

Systematic literature search across multiple databases yielded 8 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.5/10.

## 3 Results

8 papers retrieved. 15 claims extracted; 2 independently verified. Quality review score: 4.5/10.

## 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

## 5 Extracted Claims

Claim	Verified	Confidence
JaCoText is a model based on Transformers neural network designed to generate Java source code from natural language text	✓	0.35
JaCoText was initialized from powerful pre-trained models.	✓	0.17
JaCoText underwent additional pretraining on a Java dataset.	×	0.14
Experiments for JaCoText were conducted on the CONCODE dataset.	×	0.11
JaCoText-L-2CC-PL achieved a BLEU score of 39.87 on the CONCODE dataset.	×	0.05
JaCoText-L-2CC-PL achieved an Exact Match (EM) score of 22.45 on the CONCODE dataset.	×	0.04
JaCoText-L-2CC-PL achieved a CodeBLEU score of 42.49 on the CONCODE dataset.	×	0.05
JaCoText-L-2CC-PL achieved higher BLEU, EM, and CodeBLEU scores than CoText-1CC, PLBART, CodeGPT-Adp, and T5-base on the CodeSearchNet dataset.	×	0.03
CodeGPT is trained from scratch on the CodeSearchNet dataset.	×	0.02
CodeGPT-adapted is initialized from GPT-2 pretrained weights.	×	0.03
PLBART uses the same architecture as BART-base.	×	0.00
PLBART uses three noising strategies: token masking, token deletion, and token infilling.	×	0.01
CoText uses the same architecture as T5base.	×	0.00
CoText is trained on both unimodal and bimodal data using the CodeSearchNet Corpus and GitHub Repositories.	×	0.07
Reference [17] used a BiLSTM encoder and an RNN decoder to generate syntactically valid parse trees.	×	0.02

## References

- <http://arxiv.org/abs/1906.04357v3>
- <http://arxiv.org/abs/2303.12869v1>
- <http://arxiv.org/abs/1605.04464v1>