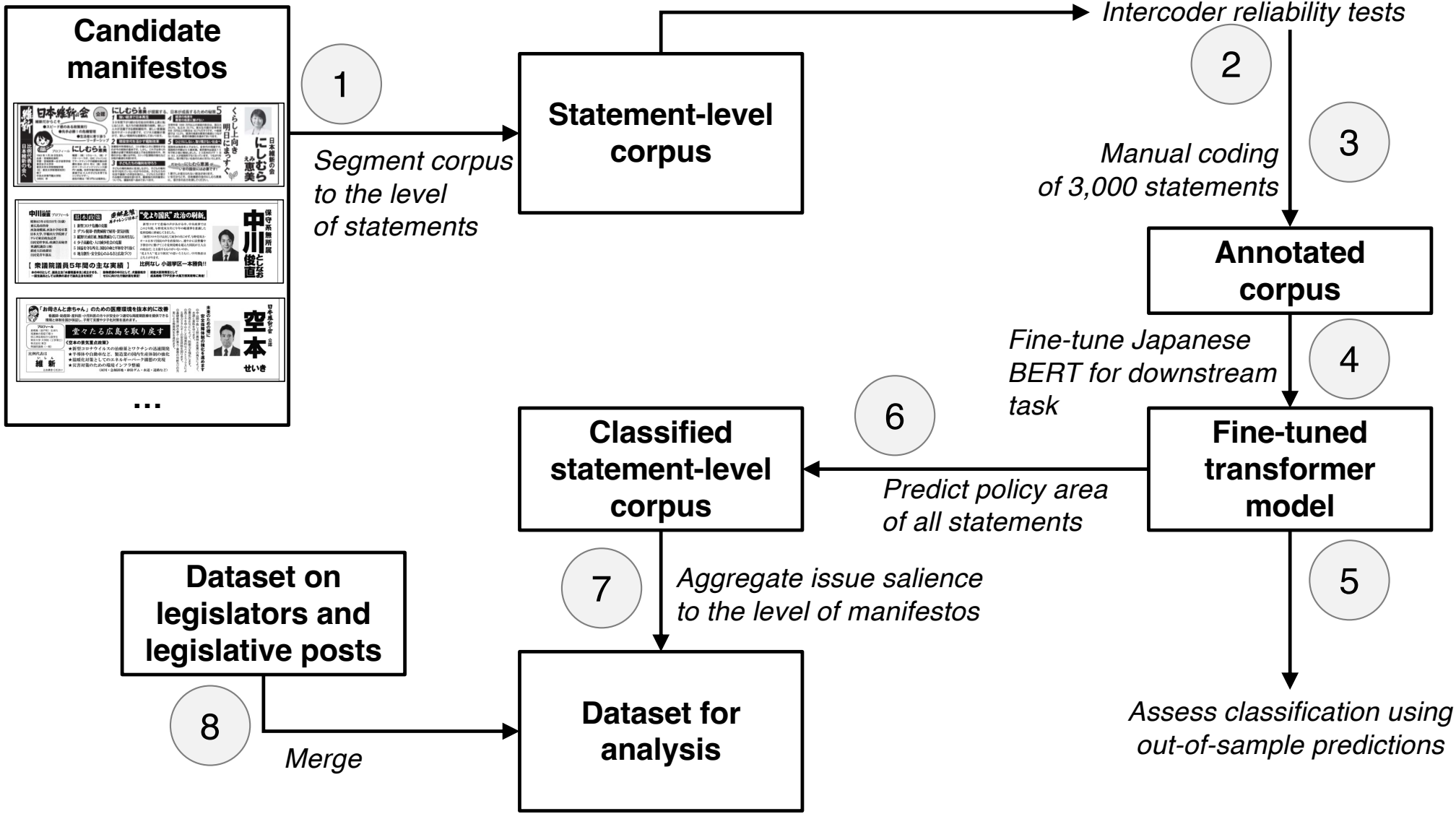
Candidate manifestos
...
1
Segment corpus to the level of statements
Statement-level corpus
Intercoder reliability tests
2
Manual coding of 3,000 statements
3
Annotated corpus
Fine-tune Japanese BERT for downstream task
4
Fine-tuned transformer model
5
Assess classification using out-of-sample predictions
6
Predict policy area of all statements
Classified statement-level corpus
7
Aggregate issue salience to the level of manifestos
Dataset on legislators and legislative posts
8
Merge
Dataset for analysis