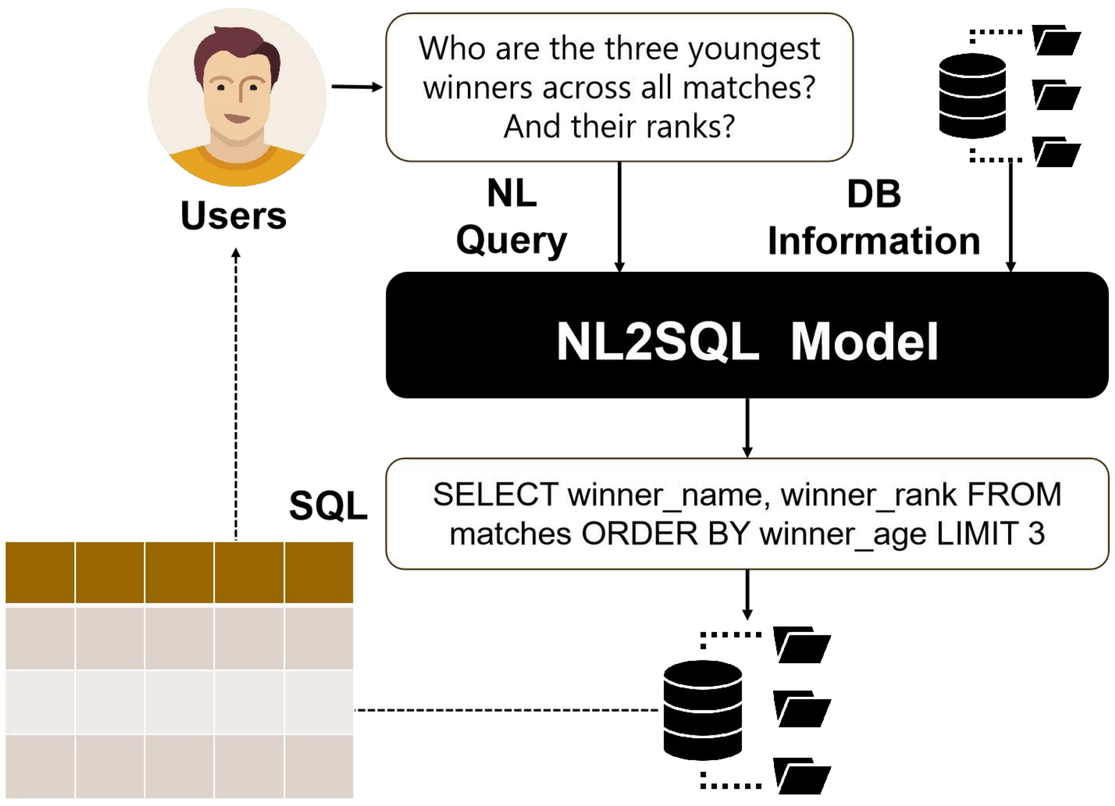
The Dawn of Natural Language to SQL: Are We Fully Ready?

Boyan Li[1], *Yuyu Luo[1,2], Chengliang Chai[3], Guoliang Li[4], Nan Tang[1,2]

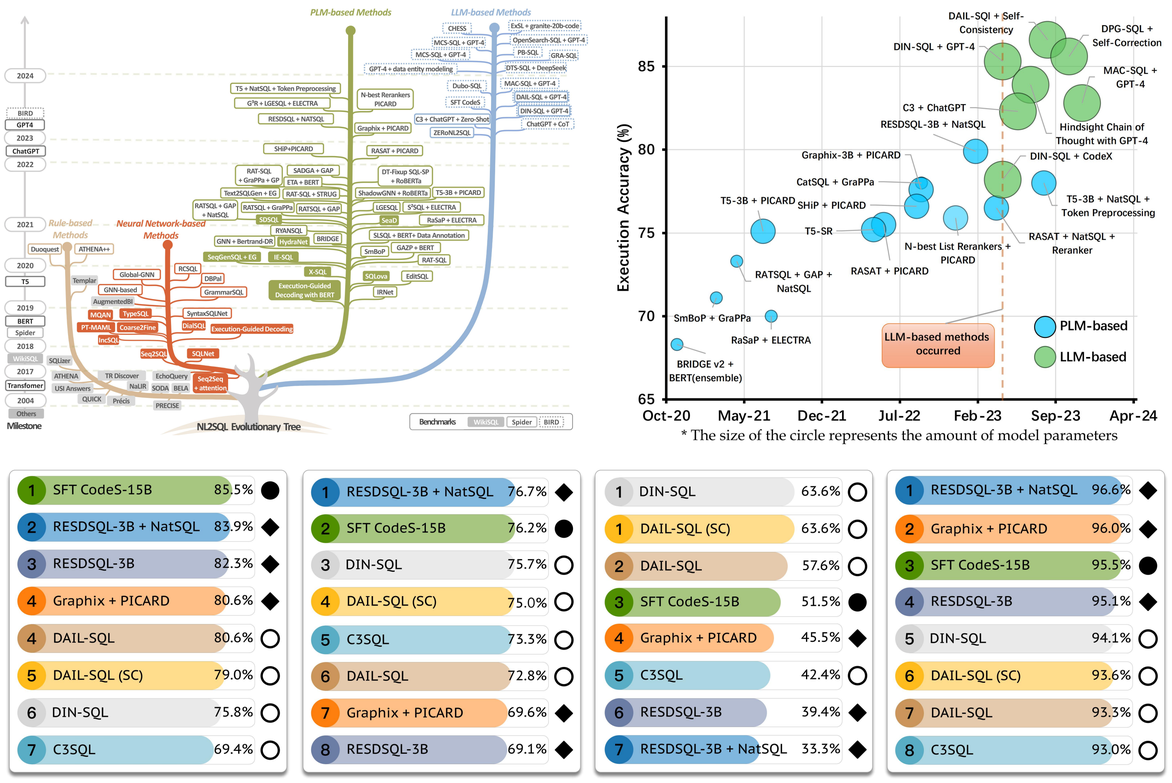
[1]HKUST(GZ), [2]HKUST, [3]Beijing Institute of Technology, [4]Tsinghua University

Yuyu Luo: yuyuluo@hkust-gz.edu.cn

Natural Language to SQL (NL2SQL)



Are We Fully Ready For Application?



(a) Test Results in Domain of *Competition* (b) Test Results in SQL with JOIN (c) Test Results in Nested SQL (d) Test Results in Query Variance

NL2SQL360: A Multi-angle Evaluation Framework

Multi-angle Evaluation

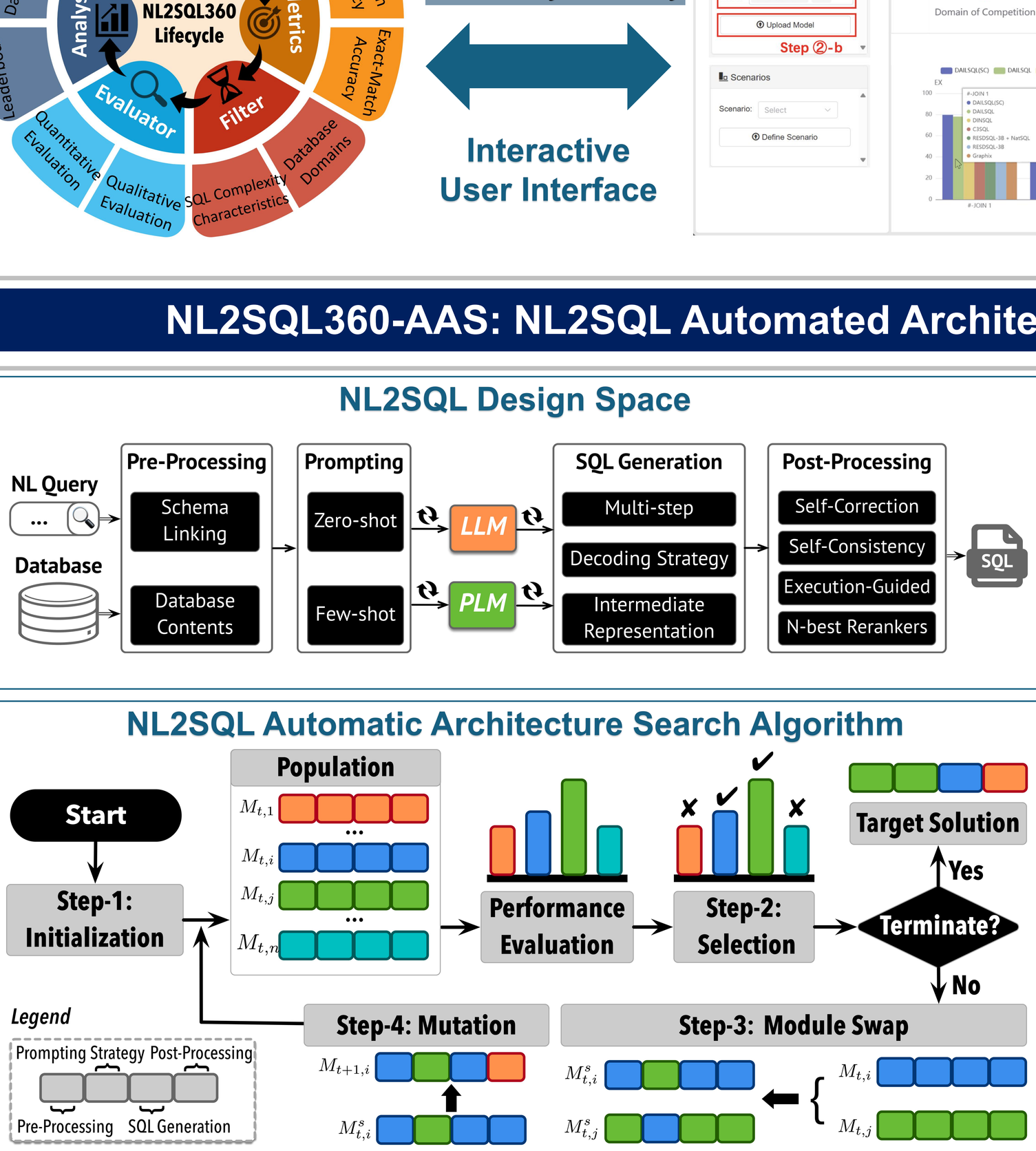
- SQL Characteristics
- SQL Complexity
- Database Domain
- NL Query Variance
- Efficiency & Economy

Interactive User Interface

NL2SQL360-AAS: NL2SQL Automated Architecture Search Algorithm

NL2SQL Design Space

NL2SQL Automatic Architecture Search Algorithm



SuperSQL

```
/* Given the following database schema (with linked
value list after column definition): */
CREATE TABLE airports (
    City text,
    AirportCode text primary key,
    AirportName text -- ["Aberdeen"]
)

CREATE TABLE airlines (
    uid int primary key,
    Airline text
)

/* Answer the following: What are airport names at City
'Aberdeen'?*/
SELECT AirportName FROM airports WHERE City = "Aberdeen";
```

Clear & Rich Prompt

Model	Spider-Dev(%)	Spider-Test(%)	BIRD-Dev(%)	BIRD-Test(%)
BRIDGE v2	70.3	-	-	-
RESDSQL-3B	81.8	-	43.9	-
C3SQL	82.0	82.3	50.2	-
DINSQL	82.8	85.3	50.7	55.9
DAILSQL(SC)	83.6	86.6	55.9	57.4
SuperSQL	**87.0** (3.4↑) (17.7↑)	**87.0** (0.4↑) (4.7↑)	**58.5** (2.6↑) (14.6↑)	**62.7** (5.3↑) (6.8↑)

Effectiveness