DATA COLLECTING

```
●──→ SELECT API ──→ SELECT SEARCH FIELD ──→ SELECT SEARCH TERM ──query──→
```

SELECT API ←─No─ Have all APIs been used? ←─Yes─ Have all fields been used? ←─Yes─ Have all terms been used? ─No→ SELECT SEARCH TERM

Project (Glynatsi, 2017)

Have all APIs been used? ⋯ stop collection ⋯→ raw data ←─ articles ← Project (Glynatsi, 2017)

CLEANING

FIND semantically similar titles
REMOVE duplicate articles

FIND semantically similar authors
FIX author names in all entries to
have the same convention

FIND articles that were collected
because a search term existed in the
text
REMOVE not relevant articles

SUPPLEMENTING

clean data   +   articles of interest   →   data set