Edinburgh Associative Thesaurus (EAT) as RDF and DBpedia Mapping

Jörn Hees, Rouven Bauer, Joachim Folz, Damian Borth, Andreas Dengel

EAT: G. Kiss, C. Armstrong, R. Milroy, J. Piper (1973)

- Association corpus
  - For each stimulus asked 100 people for a response
  - Strong responses became stimuli of next round
- 790 K raw associations (free text)
- Graph: (|V| = 23 K, |E| = 325 K)
- 5000 strong associations (>19x)
  - 167.4 K raw associations

Example Associations:

dog  cat  67
cat  dog  49
collar  neck  8
collar  tail  4
tail  neck  4
tail  collar  3

Attention:

Results:

- 1066 mapping candidates
  - Manual verification

EAT as RDF:

- 1.6 M triples

`“PUPIL”  eat:term=pupil`

`“EYE”  eat:term=eye`

Automatic Mapping Approach to DBpedia:

- Lookup + Scoring on Wikipedia

  Scoring:
  - Composite phrases ("port" - "wine")
  - Word types non-nouns ("unbound" - "free")
  - Reflexive mappings ("child" - "children")
  - Plural words ("thumbs" - "fingers")
  - Disambiguation pages ("pod" - "pea")

- Results:
  - 727 unique associations
  - between DBpedia Entities

Mapping Verification:

- Web App
  - Allows quick manual verification
  - Test group (13 people)
  - Valid: 3 people say "correct"
  - Invalid: Everything else (for now)

790 valid semantic associations
- 25.5 K raw associations