# **Obtaining Query-specific Similar Concepts with BERT-based retrieval For Commonsense Knowledge Gameboard**

#### Author: Martijn van Woensel Supervisor: Gaole He Responsible professor: Ujwal Gadiraju M.G.A.vanWoensel@student.tudelft.nl

#### 1. Background

Machine learning models require much data in order to answer questions as well as possible. **Commonsense knowledge**: knowledge every person has that is used to differentiate between objects such as different types of fruit. Most commonsense knowledge data acquisition methods are created without a specific task in mind. FindltOut[1] is a Game With a Purpose(GWAP) that obtains commonsense knowledge using a "guess who" style of game.

# 2. Main question

A query can be broken down into concepts and similar concepts and be used in GWAPS like FindItOut to obtain commonsense knowledge.

RQ: How to obtain commonsense knowledge for a specific purpose (e.g., given a query)?

- How do we find related concepts?
- How can the quality of "relatedness" be scored?
- How does it compare to a previous method used in FindltOut?

#### 3. Method

- Queries get processed into a concept.
- Using First and second-order proximity search Similar concepts can be found without being directly connected.
- The resulting concepts get scored using BERT[2] a Natural Language Processing model. Using this a selection can be made.



Figure 1. The pipeline of concept generation from a query



Figure 2. First & Second order proximity example

### 4. Results



Figure 3. Box plot of BERT score Best: highest scoring concept per question Average: average score per question

# **5.** Conclusion

Using query-specific similar concept obtainment seems to be an improvement over the previously used Greedy Max coverage algorithm in finding Similar concepts in FindItOut.



[1]. Balayn, G. He, A. Hu, J. Yang, and U. Gadiraju, "Ready player one! eliciting diverse knowledge using aconfigurable game," Virtual Event, 2022.

2]J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 201