**General Topic:** The following three papers target three perspectives to improve the performance of information retrieval tasks, including Lexical Match Consideration, Query Expansion, and Document Expansion.

**Importance:** Three following papers have shown interesting approaches for information retrieval tasks, which enable promising directions in the field. Although traditional information retrieval procedures such as BM25, a TF-IDF-based method, achieve significant results on some information retrieval tasks, it has some limitations due to vocabulary mismatch between documents and queries. However, recent work that captures dense document representations ignores the lexical match, which is also an important factor. Therefore, the following three papers address these challenges through two main approaches: introducing/generating more information to the queries and documents and using contextualized models on exact lexical matching.

**Paper 1: (Lexical Match Consideration)**

**COIL: Revisit Exact Lexical Match in Information Retrieval with Contextualized Inverted List.**

Luyu Gao, Zhuyun Dai, and Jamie Callan.


- **Specific Behavior or Activity** The specific topic of the paper is lexical matching to improve the information retrieval task.' a contextualized exact match retrieval architecture that brings semantic lexical matching.' In more detail, the paper tries to resolve the problem of soft matching. "Recent neural information retrieval approaches shift towards soft semantic matching all query document terms, but they lose the computation efficiency of exact match systems.

- **Research Questions**
  - "how much gain can we get if we introduce contextualized representations back to lexical exact match systems?"
  - "can we build a system that still performs exact query document token matching but compute matching signals with contextualized token representations instead of heuristics?"

- **Challenges** The authors consider multiple existing works in information retrieval, including BERT-reranker, Dense Retrievers, and ColBERT. However, the soft-matching system. For example, Dense Retrievers ignore the lexical match, an important factor in scoring relevance. In addition, they consider the limitation of ColBERT, which is a model that considers token-wise representations. However, through the structure of ColBERT, it is hardware engineering.

- **Paradigm** The advancements in information retrieval methodologies have enabled multiple approaches for improving information retrieval systems, including traditional methods such as BM25 to neural network-related models to consider semantic meanings of queries and documents (e.g., DPR). Therefore, the authors design and explore their models, combining the two ideas which consider lexical match and semantic meanings on token levels.

- **Claims**
  - "... exact lexical match mechanism can be greatly improved with the introduction of contextualized representation".
  - "COIL’s token-level match also yields better fine-grained signals than dense retriever’s global match signal."
  - "COIL-full further combines the lexical signals with dense CLS match, forming a system that can deal with both vocabulary and semantic mismatch, being as effective as all-to-all system."
• **State of Knowledge**
The authors show the improvements of their models compared to other baselines, including both traditional retrievers and LMs:

"Our experimental results show COIL outperforms classical lexical retrievers and state-of-the-art deep LM retrievers with similar or smaller latency."

• **Evidence**
To support their conclusions, they compare their models with other baselines on two datasets, the MS MARCO passage and MS MARCO document, on various metrics, including MRR, R@K, and NDCG@K. To evaluate the efficiency, the authors also examine the CPU and GPU utilization of the model.

• **Story Structure**
Firstly, the authors introduce two main approaches for information retrieval systems: lexical matching and soft matching. Then, they explore the limitations of two types of approaches and describe how their approaches can resolve mentioned limitations. After designing their methods, they compare the performance among all baselines and their framework to show the efficiency and correctness of their model. To show the validity of the work, they also include experiment settings and perform the analysis of different configurations of models.

---

**Paper 2: (Document Expansion)**
Soyeong Jeong, Jinheon Baek, ChaeHun Park, and Jong Park.

• **Specific Behavior or Activity**
The specific topic that the author’s study focuses on is augmenting the data, especially the documents in the task.

  - "Thus, we should address the limitations of both sparse and dense models, the vocabulary mismatch problem, and the need for a large amount of training data, respectively."
  - "In this work, we focus on document expansion and propose to abstractly generate the key information corresponding to the given document in an unsupervised manner."

• **Research Questions**
The authors do not formally define the research questions. However, the following quotes can be considered as the research questions:

  - "To expand the document for information retrieval tasks by generating document-related text, which contains novel but semantically similar terms for the given document without using query-document pairs."

• **Challenges**
The authors consider two types of models in the information retrieval representations: sparse and dense representations. The challenges are as follows: "One of the challenges in information retrieval (information retrieval) is the vocabulary mismatch problem, which happens when the terms between queries and documents are lexically different but semantically similar."

• **Paradigm**
The authors introduce their research target based on the recent advancements in information retrieval systems/techniques. Traditional and recent methods focus on learning the representations of queries and documents to achieve better performance in retrieving. Therefore, they designed the architectures using large language models to generate more information in the document to resolve the vocabulary mismatch problem of traditional models.

• **Claims**

  - It is effective when implementing the abstractive generation scheme when generating novel words for the documents.
  - "Document expansion is more effective than query expansion since a document often contains more signals than a query with its longer length."
The authors show outperformance when implementing document expansion and abstractive generation on information retrieval tasks.

Evidence To support their conclusions, they compare their models with other baselines on two datasets, ANTIQUE and MS MARCO, on various metrics, including MRR, R@K, P@K, MAP, and NDCG@K. To validate the robustness of their models, they test their models in different settings, including various language models, stochastic generation strategies, number of expanded sentences.

Story Structure Firstly, the authors introduce the challenges of the information retrieval topic. Then, they present two approaches currently being explored to resolve these issues: query expansion and document expansion. After designing their method on one of the approaches, document expansion, they compare their methods with other related work on the same task/datasets to show their improvements. Lastly, they conduct ablation to examine in depth multiple perspectives of their models and perform the qualitative and case study to examine the strengths of the introduced model.

Paper 3: (Query Expansion)
Zhi Zheng, Kai Hui, Ben He, Xianpei Han, Le Sun, and Andrew Yates.

Specific Behavior or Activity The authors’ study focuses on augmenting the data, especially the query in the task utilizing the strength of pre-trained models.

Research Questions The authors do not formally define the research questions. However, the following quotes can be considered as the research questions:

Challenges The author introduces the challenges in existing query expansion methods.

Paradigm The authors introduce their research target based on the recent advancements in information retrieval systems/methods. Traditional methods, such as PRF, utilize feedback to expand the query. Moreover, with the advancements of contextualized models on rank tasks, the paper tries to incorporate these two techniques, including the utilization of relevant feedback and contextualized models, to build a framework to improve information retrieval performance.

Claims

State of Knowledge The authors show improvement when implementing query expansion, considering relevant feedback information from documents with smaller computational cost

Evidence To support their conclusions, they compare their models with other baselines on two datasets, which are Robust04 and GOV2, on various metrics, including P@K, NDCG@K,
MAP@K To validate the robustness of their models, they test their models in different settings, including removing relevant terms from equations, replacing the techniques when computing the relevance score, and examining statistical significance relative to BERT-Large using p-value.

- **Story Structure** Firstly, the paper introduces how some existing techniques can improve information retrieval tasks, especially query expansion. Then, they explore the limitations and benefits of related work. Then, they try to incorporate different models to exploit the benefits of each model into a single framework to improve the information retrieval task. Then, they design and compare their model’s performance with other baselines on the same dataset and metrics. Lastly, they include an ablation investigation that considers the first-round reranker and how each component (equation’s term and hyperparameters) contributes to the model’s performance.