**Background & Motivation**

- **Domain Adaptation**
  - Domain Adaptation aims at adapting a general domain NMT model to a target domain.
  - Current research of domain adaptation usually considers very broad target domains.
  - The words or sentences in different sub-domains have different language phenomena.

**FGraDA Dataset**

- Wiki knowledge base
  - It contains rich monolingual resources and has additional structural knowledge.
  - We collect domain related English wikipages with the help of link relations.

- Domain Adaptation aim at adapting a general domain NMT model to a target domain.
  - There are limited time and budget to collect data (especially parallel data).
  - There are limited time and budget to collect data (especially parallel data).
  - Specific research may be needed to model fine-grained domains with other heterogeneous resources that are more available.

**Our Contribution**

- We present in-depth analyses showing that there are still challenging problems to further improve the performance with heterogeneous resources.
- We collect domain related English wikipages with the help of link relations.
- We compare different existing domain adaptation approaches and benchmark the FGraDA dataset.
- We present in-depth analyses showing that there are still challenging problems to further improve the performance with heterogeneous resources.

**FGraDA Dataset**

- We build a fine-grained domain adaptation dataset for machine translation, FGraDA, to motivate wider investigation in such scenario.
- We present in-depth analyses showing that there are still challenging problems to further improve the performance with heterogeneous resources.

**Dataset Overview**

- Adaptation resource: bilingual dictionary and wiki knowledge base
- Evaluation resource: development and test set

**Bilingual Dictionary**

- It is easier and cheaper to obtain.
- It contains domain-specific word-level correspondences between the two languages.
- We manually build a small set of bilingual dictionaries.

**Wiki knowledge base**

- It is publicly available resource.
- It contains rich monolingual resources and has additional structural knowledge.
- We collect domain related English wikipages with the help of link relations.

**Our Contribution**

- We present in-depth analyses showing that there are still challenging problems to further improve the performance with heterogeneous resources.
- We collect domain related English wikipages with the help of link relations.
- We compare different existing domain adaptation approaches and benchmark the FGraDA dataset.
- We present in-depth analyses showing that there are still challenging problems to further improve the performance with heterogeneous resources.

**FGraDA Dataset**

- We build a fine-grained domain adaptation dataset for machine translation, FGraDA, to motivate wider investigation in such scenario.
- We compare different existing domain adaptation approaches and benchmark the FGraDA dataset.
- We present in-depth analyses showing that there are still challenging problems to further improve the performance with heterogeneous resources.

**Bilingual Dictionary**

- It is easier and cheaper to obtain.
- It contains domain-specific word-level correspondences between the two languages.
- We manually build a small set of bilingual dictionaries.

**Wiki knowledge base**

- It is publicly available resource.
- It contains rich monolingual resources and has additional structural knowledge.
- We collect domain related English wikipages with the help of link relations.

**Domain Adaptation**

- Domain Adaptation aim at adapting a general domain NMT model to a target domain.
- Current research of domain adaptation usually considers very broad target domains.
- The words or sentences in different sub-domains have different language phenomena.