



CSAIL



MONIVATION

UNSUPERVISED WORD TRANSLATION

▶ Many masks in NLP rely on finding **correspondences** across domains/languages

► Ideally, solved by large amounts of parallel data. Broader applicability requires **unsupervised** approaches

▶ Word-to-word translation (a.k.a *bilingual lexical induction*) - litmus test

► Recent fully unsupervised methods perform on par with supervised counterparts [Conneau et al., 2018; Artetxe et al., 2018a,b; etc]

UNSUPERVISED WORD TRANSLATION

- ▶ Many tasks in NLP rely on finding **correspondences** across domains/languages
- ▶ Ideally, solved by large amounts of parallel data. Broader applicability requires **unsupervised** approaches
- ▶ Word-to-word translation (a.k.a *bilingual lexical induction*) - litmus test
- ▶ Recent fully unsupervised methods perform on par with supervised counterparts [Conneau et al., 2018; Artetxe et al., 2018a,b; etc]

TRANSLATION VIA WORD EMBEDDING ALIGNMENT