University of Amsterdam
Informatics Institute

The Future of Language Technology

Language Technology Lab (LTL)
ltl.science.uva.nl



Language Technology Lab Research Themes

- **From Individual Languages to Multilinguality to Language Independence:** *Current language technologies are limited to a small number of languages*
    - Develop technologies that can **generalize beyond small amounts of training data** and therefore can be applied to many more languages
    - Develop technologies that can be **trained on multiple languages** and allow to **transfer knowledge** between languages leading to better performance across multiple languages
    - Develop technologies that learn **language-independent representations** and are therefore applicable to all languages

- **Trustworthy Language Generation:** *Current language models are expressive but difficult to constrain, leading to—at best—a frustrating and unpredictable user experience and—at worst—significant harm*
    - Develop probabilistic models that are amenable to **flexible and personalized control** via both soft and hard constraints
    - Develop **uncertainty-aware probabilistic models** that know what they do not know
    - Empower language technologies with **rich contextual awareness** across modalities

Roadblocks

- **Compute power**: modern state-of-the-art deep learning models in language technology are computationally demanding, often beyond academic capabilities
- **People-centered evaluation**: Research progress is measured by benchmark performance, but do they truly reflect the needs of actual users?

The Big Picture

- **Inclusivity**: Language technologies for all, especially speakers of minority languages
- **Trustworthiness**: Language technologies that are accurate and reliable and do not generate inaccurate content