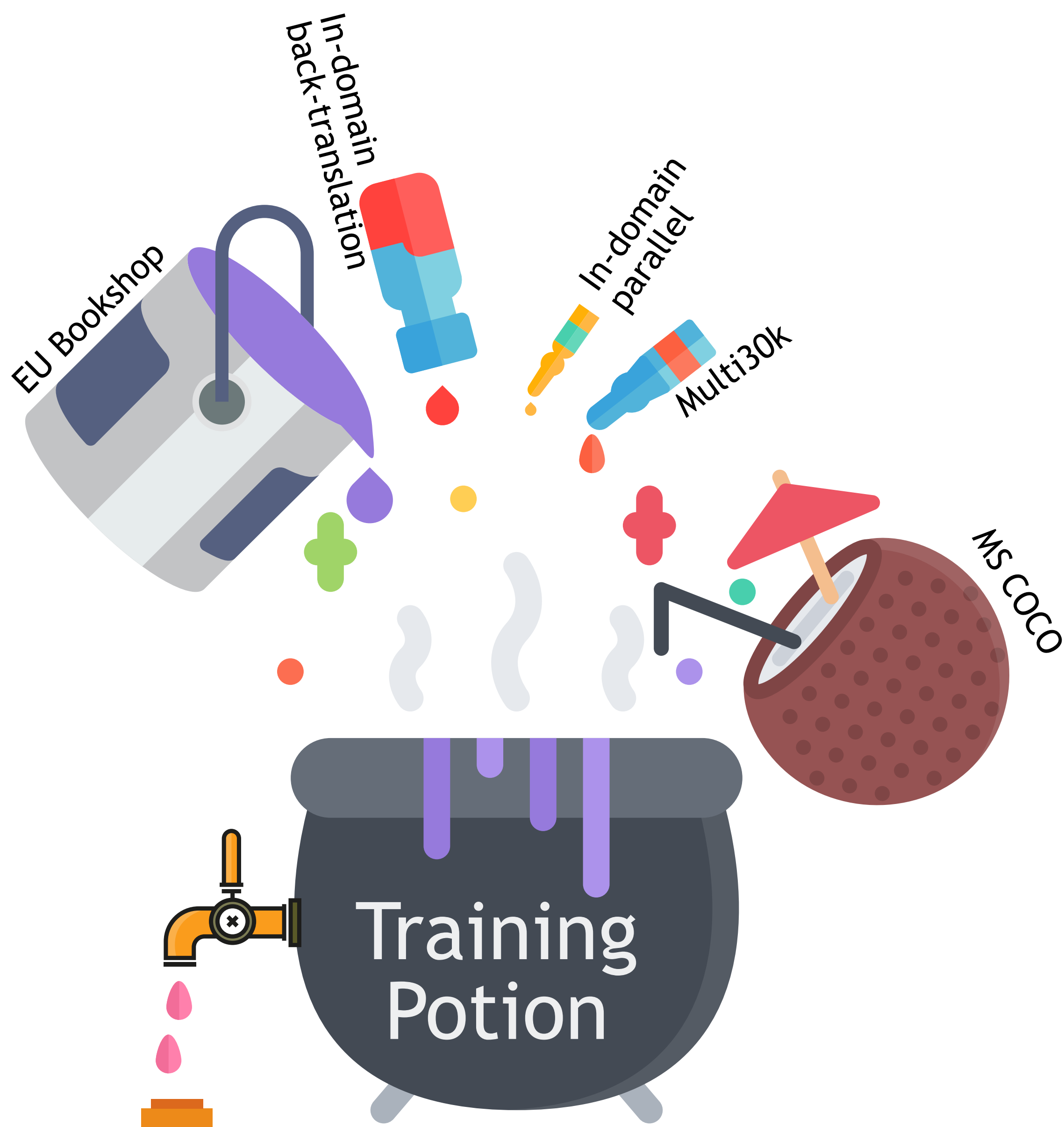




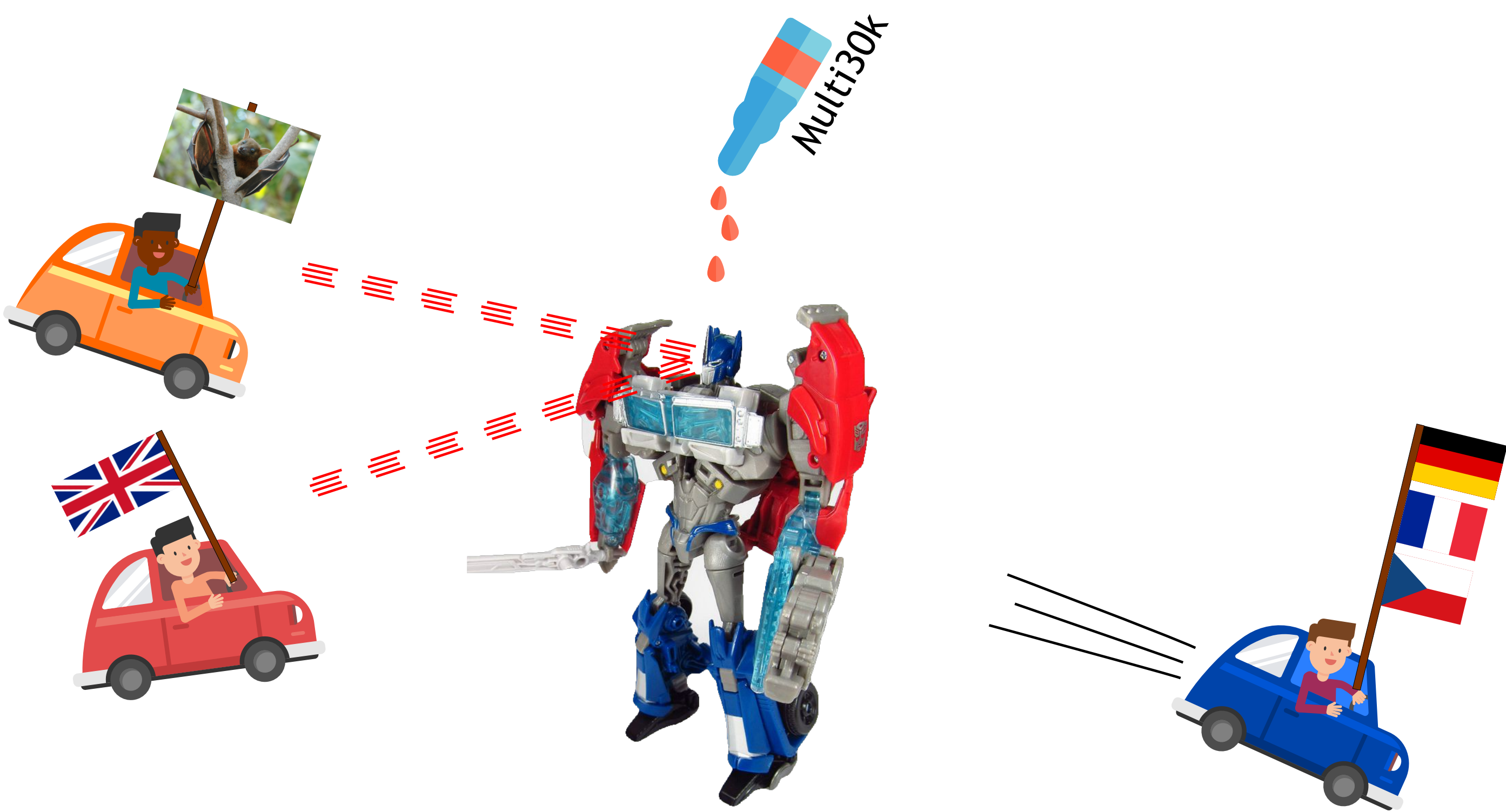
Data Cooking

- Standard dataset: Multi30k, 29k examples
- Additional in-domain data obtained by filtering large corpora using a LM
- Oversampling of under-represented data sources

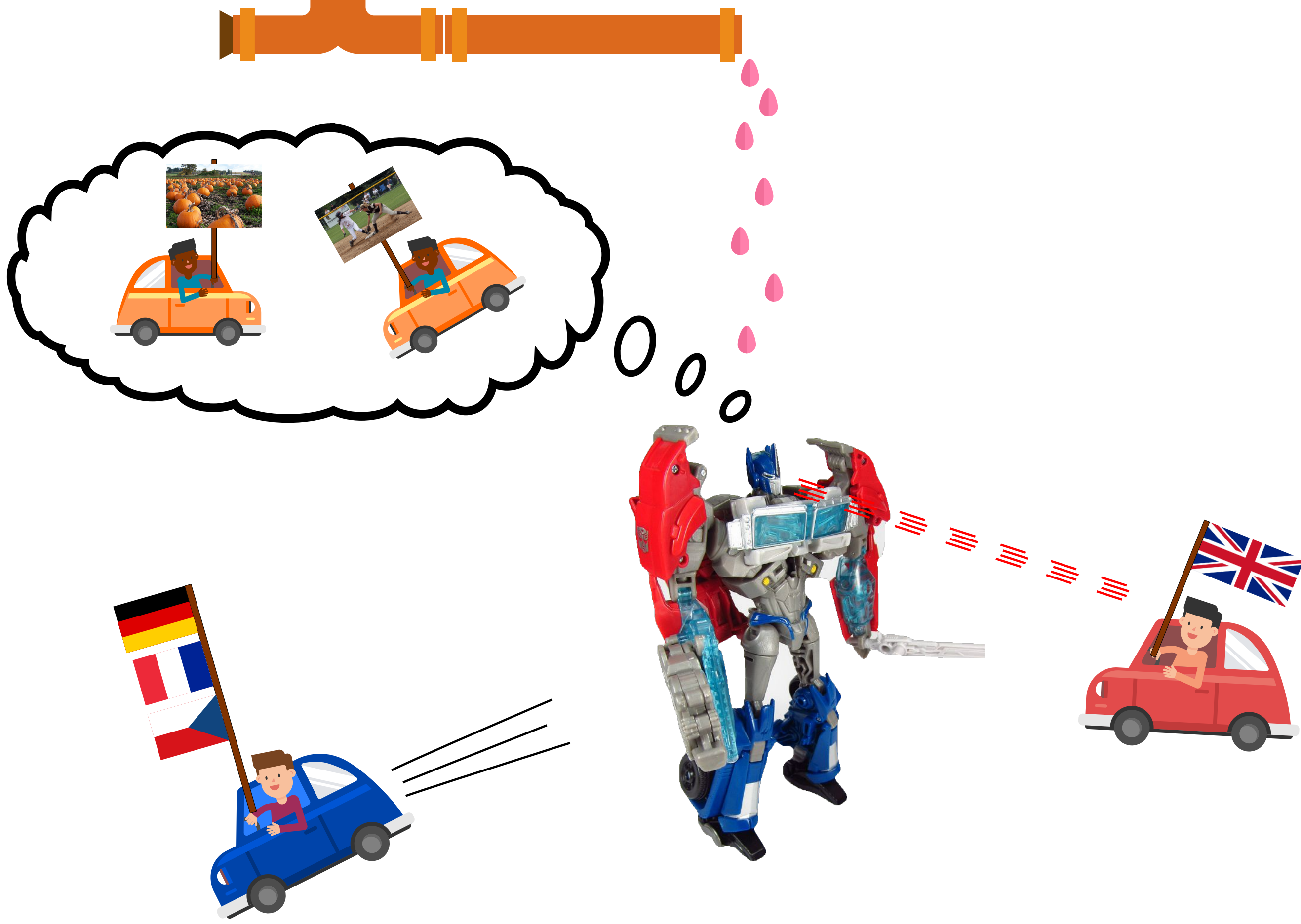
		de	fr	cs
C	Multi30k		29k	
	– oversampling factor	273×	366×	9×
	Task 2 back-translated	145k	—	—
U	in-domain parallel	3k	—	15k
	in-domain back-translated	30k	—	29k
	– oversampling factor	39×	—	7×
	EU Bookshop	9.3M	10.6M	445k
	MS COCO (English only)		414k	



Transformer Summoning



1. Doubly Attentive Transformer



2. Imagination

- Base architecture: Transformer
- Doubly Attentive Transformer: first attends to text, then to image
- Imagination: Predicting image representation as auxiliary task
- Implemented in Neural Monkey. Try it on [github.com/ufal/neuralmonkey](https://github.com/ufal/neuralmonkey)



Result Harvesting

		en-cs		en-fr		en-de	
		single	averaged	single	averaged	single	averaged
Caglayan et al., 2017		N/A	N/A	54.7 / 71.3	56.7 / 73.0	37.8 / 57.7	41.0 / 60.5
C	Textual	29.6 / 28.9	30.9 / 29.5	59.2 / 73.7	59.7 / 74.4	38.1 / 56.2	38.3 / 56.0
	Imagniation	29.8 / 29.4	30.5 / 29.6	59.4 / 74.2	59.7 / 74.4	38.8 / 56.4	39.2 / 56.8
	Multimodal	30.5 / 29.7	31.0 / 29.9	60.6 / 75.0	60.8 / 75.1	38.4 / 53.1	38.7 / 57.2
U	Textual	31.2 / 30.1	32.3 / 30.7	62.0 / 76.7	62.5 / 76.7	39.6 / 58.7	40.4 / 59.0
	Imagination	36.3 / 32.8	35.9 / 32.7	62.8 / 77.0	62.8 / 77.0	42.7 / 59.1	42.6 / 59.4