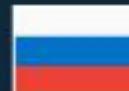


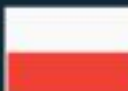


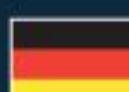







PROMT Systems for WMT 2020 Shared News Translation Task

Alexander Molchanov, Statistical and Neural MT team lead, PROMT LLC

Directions

4 language pairs, 6 directions

 **RU** ↔  **EN**  both ways
  **PL** →  **EN** 
 **DE** ↔  **EN**  both ways
  **CS** →  **EN** 

Data



Human

- WMT
- OPUS
- Private data



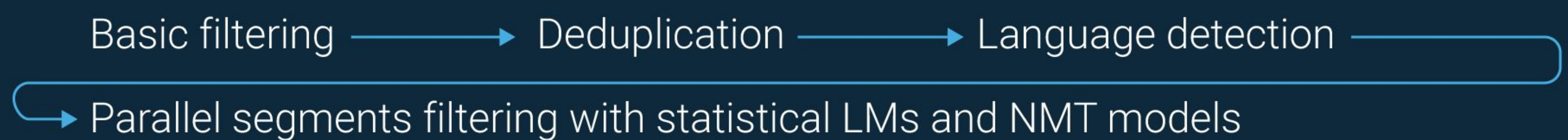
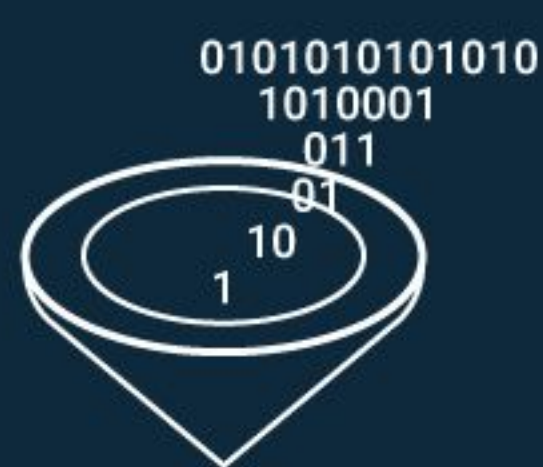
Back-Translations

- News crawl
- Wikipedia dumps
- Private data

Training Dataset Size (in human sentence pairs)

 **RU-EN**  61.1M
  **PL-EN**  37.4M
 **DE-EN**  57.9M
  **CS-EN**  40.9M

Data Filtering



Systems Architecture



- Marian NMT, baseline transformer (big transformer for EN>RU), single models
- case-insensitive BPE with case feature
 - 16K (English) + 32K (Russian) for RU-EN
 - 16K for DE-EN
 - 12K for PL-EN
 - 16K for CS-EN
- iterative back-translation for RU-EN, DE-EN
- multilingual (CS|HR|SK|SR) system for CS-EN

Results

BLEU scores for newstest2020

	best submitted system	PROMT NMT		best submitted system	PROMT NMT
 EN-RU 	25.5	23.2	 EN-DE 	41.6	34.2
 RU-EN 	55.2	53.5	 DE-EN 	58.9	53.9
 CS-EN 	29.9	25.6	 PL-EN 	34.4	31.2