

Recent Performance of NMT on News

Seg-Level English→Czech 2018

| | Ave. % | Ave. z | System |
|---|--------|--------|---------------------------------|
| 1 | 84.4 | 0.667 | CUNI-Transformer |
| 2 | 79.8 | 0.521 | UEDIN |
| | 78.6 | 0.483 | Professional Translation |
| 4 | 68.1 | 0.128 | ONLINE-B |
| 5 | 59.4 | -0.178 | ONLINE-A |
| 6 | 54.1 | -0.354 | ONLINE-G |

Doc-Aware English→German 2019

| Ave. | Ave. z | System |
|------------------------------|--------|---------------------------------|
| 90.3 | 0.347 | Facebook-FAIR |
| 93.0 | 0.311 | Microsoft-WMT19-sent-doc |
| 92.6 | 0.296 | Microsoft-WMT19-doc-level |
| 90.3 | 0.240 | Professional Translation |
| 87.6 | 0.214 | MSRA-MADL |
| 88.7 | 0.213 | UCAM |
| 89.6 | 0.208 | NEU |
| 87.5 | 0.189 | MLLP-UPV |
| 87.5 | 0.130 | eTranslation |
| 86.8 | 0.119 | dfki-nmt |
| 84.2 | 0.094 | online-B |
| ... 10 more systems here ... | | |
| 76.3 | -0.400 | online-X |
| 43.3 | -1.769 | en-de-task |

See lecture #1 for all caveats of MT evaluation.