ELG Pilot Project:
Open Translation Models, Tools and Services
Available software:

- MT server solution based on Marian NMT
- dockerised web-app, translation interface and API
- NMT training pipeline (OPUS-MT-train)
- CAT integration (OPUS-CAT)

Pre-trained translation models:

- number of bilingual models: 1,048
- number of multilingual models: 53
- number of supported source languages: 229
- number of supported target languages: 222
- number of supported language pairs: 1,715
OPUS-MT – Public NMT Models & Tools
https://github.com/Helsinki-NLP/Opus-MT

Development demo:
https://translate.ling.helsinki.fi/ui/sami

OPUS-CAT - plugins and local MT engines
https://github.com/Helsinki-NLP/OPUS-CAT

{  'alignment': [  '0-0 0-2 1-1 2-3',  '0-0 1-1 3-2 4-3 5-4'  ],  'result': 'How are you? The translation is fun.',  'server': '192.168.1.18:20001',  'source': 'fi',  'source-segments': [  'Mitä kuuluu?',  'Kuolee on hauskaa.'  ],  'source-sentences': [  'Mitä kuuluu?',  'Kuolee on hauskaa.'  ],  'target': 'en',  'target-segments': [  'How are you?',  'The translation is fun.'  ],  'target-sentences': [  'How are you?',  'The translation is fun.'  ]}

OPUS-MT at huggingface
https://huggingface.co/Helsinki-NLP

Model: Helsinki-NLP/opus-mt-ROMANCE-en

How to use this model directly from the transformers library:
from transformers import AutoTokenizer, AutoModelForSeq2SeqLM

tokenizer = AutoTokenizer.from_pretrained("Helsinki-NLP/opus-mt-ROMANCE-en")
model = AutoModelForSeq2SeqLM.from_pretrained("Helsinki-NLP/opus-mt-ROMANCE-en")

List all files in model - see also config file