

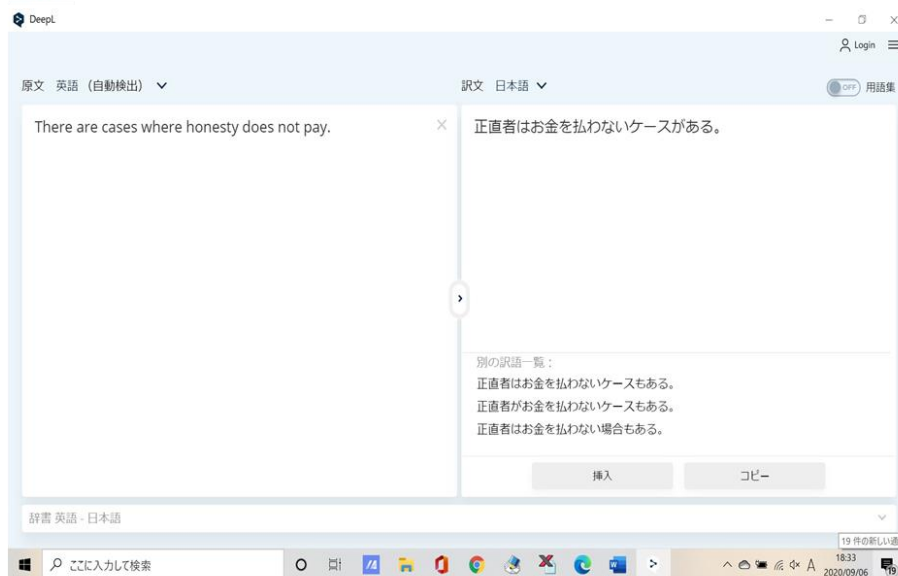
インバウンドコラム

機械翻訳と SEO

進化と限界

GPT-3 (Generative Pre-trained Transformer 3)は、「1750 億のパラメータ(変数)を持つ自己回帰型言語モデル」で、機械学習のために、Web サイトから数年にわたり収集した 1 兆語におよぶデータセットを使用し、英単語や短い文章をインプットすると、関連する「それらしい」テキストを自動生成するツールです。(参照※1)登場以来、もしかしたら、Google の SEO でも使えるのではないかと話題になっていましたが、Google の Gary Illyes 氏がツイッターで、「機械翻訳されたコンテンツはインデックスしない」と投稿しました。(参照※2)ディープラーニングの機械翻訳は、確かに日進月歩ですが、まだまだ未完成な技術です。例えば、「Honesty doesn't pay.(正直者は馬鹿を見る)」ということわざですが、精度が高いと評判の DeepL で翻訳してみると、成句であればきちんと翻訳してくれますが、否定形を短縮しない形にするなど、少しでも変わると、「正直者はお金を払わない」と逆の意味に翻訳してしまいます。(参照※3、図 1)さらに、いかにも「それらしい」日本語訳を出力するので、誤訳や訳漏れにとても気づきにくく、翻訳者界限では、「絶対にこんな校正したくない」と戦々恐々としています。

図 1



出典元: @TNK_KNCH on Twitter

※1 新 AI ツール - ITmedia ビジネスオンライン

<https://www.itmedia.co.jp/business/articles/2007/29/news025.html>

※2 機械翻訳は NG - SEO Round Table

<https://www.seroundtable.com/google-machine-translated-content-29957.html>

※3 DeepL の実力を見よ | @TNK_KNCH

https://twitter.com/TNK_KNCH/status/1302540714731945984?s=20