

日本語テキストデータ「自然言語処理」
(発注イメージ)

[で]*{NOISE}あと*{parsing|パーズング}で書いてますけど、{parsing|パーズング}ていうのは[あの]構文解析とか統語解析とか日本語では言われますけど[あの]文の構造を解析したいと[で]構造で何かていうのは、あとで[あの]ちょっと例を示したりします。

[で]中村さんは[あの][ま]ここの[えっと]客員講座の[えと]准教授の方で、[えっと]本務は<NAIST|ナイス>ていう[あの]奈良先端科学技術大学院大学の方です。

[で][えっと]彼は[あの]機械翻訳がメインで、ずっと古くから、古くからって変ですけど、前から機械翻訳をやられているので、[えっと][ま]機械翻訳ていうのは、我々言語処理の最も最初の応用で、ていうかむしろ[あの]コンピューターを最初に作られた時ていうのは、もちろん数値計算を目的にして、目的にして作られたんですけど[あの]数字だけではなくて文字データも扱えるということで、文字データを扱えるんなら[あの]翻訳ができるんじゃないかということで、[あの]最も最初に考えられた応用で、[えっと]もう[あの]*{NOISE}コンピューターができて十年も経たないうちに、そういうことが議論されている*という[ま]言ってみれば初めての数値計算以外の応用ということで、[あの]非常に歴史が長いんですけど[あの][ま][あの]何をやるかは非常に簡単で[えっと]ある言語で書かれた文ていうか[||き]記述を、[えっと]他の言語に[ま]翻訳すると[ま]結局意味を保持したまま翻訳するということですけど、[ま]これが[あの][||もと]一番最初のアイデアとしては[あの]暗号化された言語を元に戻すというような概念で、ロシア語を英語に直すというようなことができるんじゃないかというようなことが[ま]言われて[で]その後[ま][あの]いろんな研究が行われてますので、中村さんが[ま]その歴史とか[ま]最近の{algorithm|アルゴリズム}について話をしてくれます。

上記の発注文字数：695文字

*発注文字数は、マイクロソフトWord2016Windows版の「文字アカウント」機能を用いて導き出した「全角文字+半角カタカナの数」の数値とする。