

【資料2】

日本語テキストデータ (発注イメージ)

あと parsing して書いてますけど、parsing というのは構文解析とか統語解析とか日本語では言われますけど文の構造を解析したいと構造で何かっていうのは、あとでちょっと例を示したりします。

中村さんはここの客員講座の准教授の方で、本務は NAIST という奈良先端科学技術大学院大学の方です。

彼は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、機械翻訳っていうのは、我々言語処理の最も最初の応用で、ていうかむしろコンピューターを最初に作られた時っていうのは、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翻訳ができるんじゃないかということで、最も最初に考えられた応用で、もう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、非常に歴史が長いんですけど何をやるかは非常に簡単である言語で書かれた文ていうか記述を、他の言語に翻訳すると結局意味を保持したまま翻訳するというんですけど、これが一番最初のアイデアとしては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言われてその後いろんな研究が行われてますので、中村さんがその歴史とか最近の algorithm について話をしてくれます。

上記の発注文字数：599 文字

*発注文字数は、マイクロソフト Word 2013 Windows 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。