

山田 篤教授の模擬授業

情報

音声認識の仕組み

～コンピュータはどうやって「雨」と「飴」を聞き分けているの？～

講師:山田 篤

テキストを入力しなくても、スマートフォンやスマートスピーカーに話しかけると自動的に検索をしたり、家電を動かしてくれたりするようになってきています。一方で、日本語には「雨」と「飴」のように同じ発音で異なる意味を持ついわゆる同音異義語がたくさんあります。コンピュータはこれらをどうやって区別しているのでしょうか。日本語は基本的に高低アクセントであり、人間が発音するとき、「飴」と「雨」ではアクセントが異なります。この情報を使えば区別できるようになるのでしょうか。しかし、困ったことに、たとえば東京の人と関西の人とではこのアクセントが異なりますし、更には関西地方でも大阪の人と丹波の人とでは異なる場合があります。でも、住む地域毎にアプリの設定を変えるなどといった話は聞いたことがないと思います。この授業ではこれらの疑問について考えながら、一般的な音声認識の仕組みを解き明かしていきます。

■ 受講人数の目安:5～40人程度

■ 所要時間の目安:50分

■ 高校でご準備いただきたいもの:プロジェクト

情報

文字の世界

～日本で使われている漢字は何文字あるの？～

講師:山田 篤

漢字は日本だけではなく、中国を始めいくつかの国で使われていて、国毎に違った形の漢字を使っています。ここでは日本国内に限定して、日本語で使われる漢字について考えてみましょう。さて、日本語で使われる漢字の種類は何文字くらいあるのでしょうか。コンピュータの世界で通常使われる文字は日本工業規格 (JIS) という規格によって定められていて、現在第1水準から第4水準まであります。一方で戸籍などの人名用に使われる漢字は歴史的な経緯もあって非常に多様です。実は戸籍がコンピュータによって取り扱われるようになったのは比較的最近のことで、長い間ずっと手書きで運用されていたのです。皆さんの書く手書きの文字も一人一人形が違ってきますね。では、どうやって形が異なる文字を同じ文字だと見なすことができるのでしょうか。この授業ではこれらの疑問について考えながら、コンピュータでの文字の取扱い方について見ていきます。

■ 受講人数の目安:5～40人程度

■ 所要時間の目安:50分

■ 高校でご準備いただきたいもの:プロジェクト