

鹿児島大学数理情報科学談話会

第223回

日時：2017年6月15日（木）16:30 - 17:30

場所：理学部2号館502室

講師：佐藤 健一氏（広島大学）

題目：被ばく・被災関連資料に対する統計的テキスト解析の取り組み

Abstract:

原爆投下から70年が過ぎ生存する原爆被爆者の平均年齢は80歳を超える中、被爆体験などをどのように次世代に伝えていくかが社会的な問題として取り上げられるようになってきている。

一方で、広島には被爆体験、当時の手記など読み切れないほどの膨大なテキスト情報が当研究所をはじめ、国立広島原爆死没者追悼平和祈念館、広島平和記念資料館など複数の公的機関に保管されており、その活用が望まれている。

また、福島第一原子力発電所事故から6年が経過してもなお、福島には放射線への社会不安が残っている。放射線被ばく不安の要因を調査することを目的とした従来のアンケートはある程度の回答を想定して作成されるため、多様な要因を拾い上げたい場合には十分ではない。そこで、少人数グループによる聞き取り調査が試みられているが、音声データをテキスト情報に変換するとかなりの分量になり、これを定量的に把握するための解析手法が望まれている。

そこで、これらのテキスト情報を統計的な手法によって定量的に要約することを試みる。要約により、全体像の把握が容易になるだけでなく、単に読むだけでは認識することが難しい、例えば、時間変化などについても定量的に検討できるようになり、客観的かつ端的な理解につながる可能性がある。また、原文が難しい言葉を含んでいてもキーワードとしてまとめることで、子供であってもある程度は内容を想像できるなど、結果として、人々が内容に触れ易くなり、テキスト情報の活用の幅が広がると考えられる。

本報告ではテキスト解析の基本的なアプローチを紹介し、その応用例として、1) 毎年、広島市長によって読み上げられる広島市の平和宣言、2) 福島での聞き取り調査データ、取り上げ、時間軸方向の変化に着目したアニメーションなどを紹介する。

お問合せ：談話会委員 田中 恵理子 (✉ erico@sci.kagoshima-u.ac.jp ☎ 099-285-8988)