

平成27年度 情報工学コース卒業研究報告要旨

| | | |
|---|---------------------|-------|
| 外山研究室 研究室 | 氏 名 | 永 井 輝 |
| 卒業研究題目 | 法令文書スキーマの自動変換に関する研究 | |
| <p>法情報の共有や二次利用を国際的に促進するためには、単に法令文書を提供するだけでなく、国際標準規格の文書スキーマに基づいて構造化した上で法令文書を提供することが求められる。Akoma Ntoso は、法令・議会情報等を記述するための XML スキーマであり、2012年に国際標準化団体 OASIS によって採択された。Akoma Ntoso はさまざまな国や組織で用いられる法令文書を構造化できるよう柔軟に設計されており、実際にケニア、イギリス、香港などさまざまな国、地域で Akoma Ntoso に対応する取り組みが進められている。</p> <p>日本では、Akoma Ntoso に基づいた法令文書の提供はされていないが、河地らの研究で日本法令を Akoma Ntoso に準拠して構造化するための方法論が提案されている。河地らは、Akoma Ntoso が構造の柔軟性を持つために、一つの構造に対するマークアップを一意に決定できない曖昧性の問題を指摘し、日本法令の要素構造に特化した XML スキーマ (以下、日本法令スキーマ) を活用したアプローチを提案している。日本法令スキーマは日本法令外国語訳データベースシステム (以下、JLT) で採用されており、日本法令の文書に対して、曖昧性のない一貫した構造化が可能である。日本法令スキーマから Akoma Ntoso 形式に変換することで、Akoma Ntoso の国際共有性と、日本法令スキーマの厳密性を両立できることを示し、実際に変換規則を定義した。しかし、数例の日本法令文書に対して Akoma Ntoso 形式への変換が手動で行われたにすぎず、自動変換器の開発や定量的な評価は行われていない。</p> <p>そこで、本研究では、日本法令スキーマに準拠した XML 文書から Akoma Ntoso に準拠した XML 文書を生成する自動変換器を開発する。本変換器では、まず、JLT において公開されている日本法令スキーマ準拠の XML 文書を Ruby の XML パーサである REXML を使用して解析する。次に、河地らが作成した変換規則に基づいて Akoma Ntoso 形式の XML 文書に自動変換する。</p> <p>開発した自動変換器を定量的に評価するため、2015年5月時点で JLT から提供されている 638 法令分の XML 文書 (日本語 638 文書、英訳 638 文書、合計 1,276 文書) を Akoma Ntoso 形式の XML 文書に自動変換する実験を行った。変換された文書に対する評価として、XML 文書の妥当性検証機能を持つ xmllint を用いて、Akoma Ntoso 形式で妥当に変換されている文書数を計測した。その結果、日本語 624 文書 (97.8%)、英訳 618 文書 (96.9%)、全体として 1,242 文書 (97.3%) で文書構造が妥当であることが確認できた。妥当でないと判断された 34 文書 (日本語 14 文書、英訳 20 文書) に対して原因を調査したところ、JLT から提供されている文書において、日本法令スキーマに対しては妥当であるが、元の法令と対応した正しい構造化が行われていないなどの問題があることが判明した。そこで、34 文書の構造を修正した上で再実験を行ったところ、すべての文書において妥当に変換できたことを確認した。</p> <p>この実験結果から、開発した自動変換器は、日本法令スキーマに従って正しく構造化されている XML 文書に対しては、Akoma Ntoso に準拠した XML 文書に自動的に変換できることを確認した。よって、本研究により、日本法令を国際共有に適した Akoma Ntoso 形式でいつでも提供できるようになった。</p> | | |