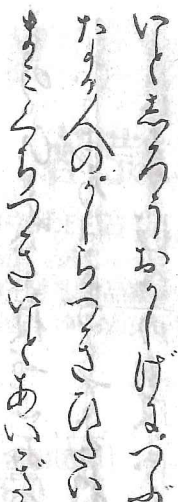


凸版印刷は江戸時代までに書かれた書物や古文書の「くずし字」を写真は源氏物語、国文学研究資料館所蔵書を自動で判読し、電子テキストデータに置き換える技術を開発した。8割以上の精度で判読することができるという。今夏から試験的な解読サービスを始め、2016年度中に大学や博物館などに本格的に売り込む。

くずし字は楷書とは異なり簡略化した文字を連ねていくため、一文字の

自動判読精度8割

古文書の「くずし字」



みでは判読できないこと装置(OCR)で書物の多い。凸版印刷の技術字を読み込み、自動判読では、まず、くずし字を誰でも読める文字に置き換えるための解読用データベースを作成。そのうえで光学式文字読み取り結果をパソコン画面に表示

凸版がデータ化技術

示す。サービス価格は1冊あたり2千円からを想定する。

凸版印刷によると、最近では専門家の減少や資料の経年劣化などを背景に、歴史的資料の電子化のニーズが高まっているという。江戸時代以前のくずし字で書かれた書籍は100万点以上とされ、テキスト化は、専門家が実際に書物を読みながら文字を判読し手作業でパソコンなどに打ち込んでいくケースが多い。