

# Digital Bangla text



EARLIER if you wanted to digitise a magazine article or a book you had to spend hours typing the content and correcting the typing mistakes. You can use a scanner or a digital camera to create an image or a snapshot of the document but the content cannot be edited.

Now, Optical Character Recognition (OCR) software has made it possible to convert any written, printed or image character into machine-readable text files which can be edited and searched very easily.

Although OCRs have been available for English text for quite some time, there was no such software for Bangla. But recently, United international University's (UIU) CSE Department developed the first-ever Bangla OCR software. The research project was granted Tk 2.3 million from the service innovation fund of a2i programme. The OCR will enable digital archiving of new and old Bangla publications. It will also save a lot of time and effort when creating, processing and repurposing various documents.

