

Die Aufgabe der Registerstelle (RST) des Epidemiologischen Krebsregisters Niedersachsen besteht u. a. in der Erfassung und Bearbeitung von Tumordaten. Für eine Optimierung einzelner Arbeitsschritte ist es wichtig Arbeitsabläufe effizienter zu gestalten. Ein Beispiel dafür ist das Verfahren, Befundkopien einzuscannen.



## Einführung

Im Krebsregister Niedersachsen gehen derzeit pro Quartal ca. 25.000 Pathologiemeldungen ein.

Der Großteil dieser Meldungen (80 %) wird über den elektronischen Weg geliefert und im Rahmen der Qualitätssicherung weiter bearbeitet. Ca. 5000 Meldungen (20%) gehen in Form von Papiermeldebögen oder maschinenlesbaren Befundkopien ein (s. Abb. 1).

Auf den handschriftlich ausgefüllten Papiermeldebögen sind nur die wichtigsten Angaben wie Diagnosedatum, Lokalisation, Seite, Histologie, Dignität, Grading, TNM-Werte und eine kurze Angabe zum Klartext enthalten. Die Eingabe dieser Bögen erfolgt manuell.

Die Befundkopien enthalten entsprechend des Textgehaltes weitaus mehr und ausführlichere Informationen, die für die weitere Bearbeitung bzw. Kodierung von großer Bedeutung sein können.

## Methoden

Um die Dateneingabe und die Qualitätssicherung dieser Befundkopien, bei Paketgrößen von bis zu 500 Meldungen, effizienter und schneller zu gestalten, arbeitet die Registerstelle seit 2003 mit der Scannertechnologie.

Hierzu werden die vorliegenden Diagnosetexte mittels eines Scanners und der OCR-Software „Abyy FineReader 6.0 Corporate Edition“ automatisch über einen Einzelblatteinzug erfasst. Einer Kontrolle bedarf es hier nur selten, da Zwischenfälle wie ein Papierstau kaum vorkommen. Nach dem Scannen liegen die Dateien in drei Varianten vor: als Original auf Papier, als Bilddatei und als Textdatei, welche für den Import benötigt wird. Die Bilddatei entspricht dabei dem Original und bleibt unverändert. In ihr werden lediglich die von der OCR-Software nicht eindeutig erkannten Buchstaben farblich hinterlegt (s. Abb. 2).

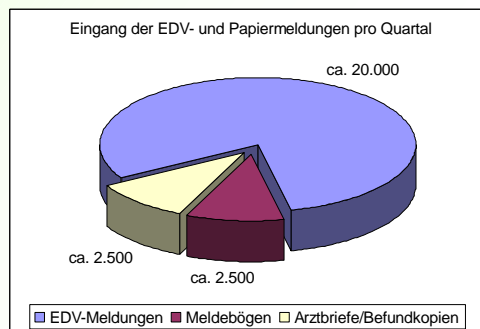


Abb.1: Dateneingang pro Quartal

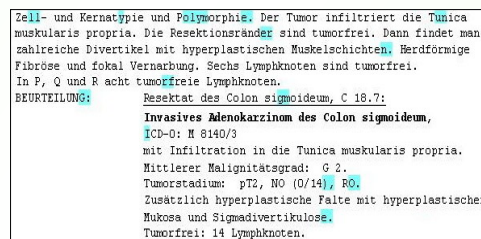


Abb.2: Scan des Arztbriefes mit Anmerkungen der OCR- Software zu unklaren Buchstaben

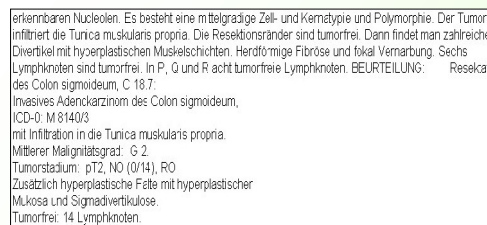


Abb.3: Textdatei, welche für den Import verwendet wird

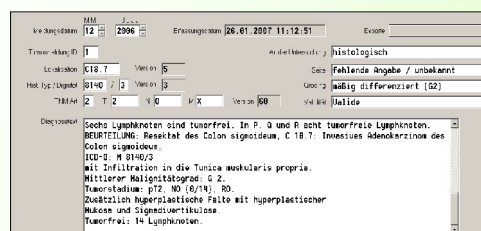


Abb. 4: Bearbeitungsmaske CARTRUST, mit beim Import automatisch eingefügten Kodierungen

Die bei dem Scan entstandenen Textdateien werden über das für die RST entwickelte Programm CAMEL in die Datenbank importiert (s. Abb. 3).

Dabei werden diese sogenannten OCR-Dateien zuerst mit einem Modul in xml-Dateien umgewandelt und anschließend mit Hilfe eines entsprechenden Skriptes (XSLT-Trafo) in die Datenbank eingelesen. Die einzelnen Befunddaten liegen nun in der dafür vorgesehenen Bearbeitungsmaske CARTRUST vor (s. Abb. 4). Erst hier werden die möglichen Fehler, z.B. durch eine Fehlinterpretation einzelner Buchstaben oder selten durch die fehlerhafte Worterkennung aufgrund zusätzlicher handschriftlicher Vermerke, sichtbar. Der Bearbeiter hat allerdings nun die Möglichkeit die zwei anderen Medien zur Kontrolle heranzuziehen.

Sofern die Programmierung der Software und die Art der Befundtexte es zulassen, können Angaben wie z.B. Grading, Lokalisation, Histologie, TNM und Datumsangaben beim Import automatisch übernommen werden. Dies ist bei etwa 50 % der zu scannenden Befunde möglich.

## Fazit

Die Scannertechnologie wird in der Registerstelle des EKN seit 2003 erfolgreich eingesetzt, da die manuelle Erfassung der Befunde entfällt und Arbeitsprozesse parallel ablaufen können. Dadurch reduziert sich der Arbeits- sowie Zeitaufwand erheblich.

Die Übernahme des gesamten Textinhaltes in die Datenbank führt zusätzlich zu einer verbesserten Datenqualität.

Die Kosten belaufen sich auf die Anschaffung der Hardware (PC, Scanner), sowie die Programmierung und Pflege der Software.