

## Digitales Diktieren und elektronische Spracherkennung – 2004

Die Einsatzmöglichkeiten von Spracherkennungssystemen werden in der Justiz seit mehr als 10 Jahren mit großen Erwartungen untersucht. Die Erörterung und Diskussion von Vor- und Nachteilen der Online- und Offline-Spracherkennung sowie verschiedenen Diktiersystemen war Gegenstand des Arbeitskreises „Digitales Diktieren und elektronische Spracherkennung – 2004“ beim 13. EDV-Gerichtstag<sup>1</sup>. Als Referenten waren Herr Richter am Oberlandesgericht Karl-Heinz Volesky (OLG Hamm) und Herr Präsident des Landgerichts Wolfgang Prahl (LG Detmold) erschienen. Die folgenden Ausführungen sollen die Ergebnisse des Arbeitskreises wiedergeben.

### 1.) Online- und Offline-Spracherkennung

#### a) Worum geht's?

Bei der Spracherkennung gibt es zwei verschiedene Arbeitsmöglichkeiten: Online- und Offline-Spracherkennung. Bei der **Online-Spracherkennung** wird das Diktat über ein Mikrofon am PC diktiert. Der Text erscheint – leicht verzögert – direkt am Monitor, sodass Korrektur, Weiterverarbeitung und Speicherung unmittelbar erfolgen können. Bei der **Offline-Spracherkennung** wird das Diktat zunächst auf einem digitalen Aufnahmegerät/Datenträger aufgenommen und als Audiodatei im DSS-Format (Digital Speech Standard) gespeichert. (DSS-Dateien zeichnen sich dadurch aus, dass sie bei spracherkennungstauglicher Aufzeichnungsqualität zugleich nur geringe Speicherkapazitäten beanspruchen.) Erst nach Abschluss des Diktats wird die Audiodatei an die Spracherkennungssoftware weitergeleitet. Diesen Arbeitsschritt sowie die weitere Bearbeitung kann eine Schreibkraft übernehmen.

#### b) Erfahrungen im Überblick

Die Möglichkeiten der Spracherkennung werden im Bereich der Justiz seit 1995 beobachtet. Seit 1998 werden zusätzlich vereinzelt Praxistests durchgeführt. Der erste praktische Einsatz von Spracherkennung findet in einigen Bundesländern ab 2001/2002 statt. Ergebnis: Die Online-Spracherkennung ist für die überwiegende Justizpraxis kein Konzept. Zwar sind die Erkennungsergebnisse meist gut brauchbar, dagegen spricht aber, dass die Handhabung nicht justiztypischen Arbeitsabläufen entspricht. Es bedarf also eines neuen organisationsverträglichen Lösungsansatzes. Hier bewährt sich die Offline-Spracherkennung: Autor und

---

<sup>1</sup> Im Internet: <https://www.edvgt.de/Tagung04/ak04/spracherkennung.shtml>

Schreibkraft können örtlich und zeitlich verschieden das Diktat bearbeiten. Je nach Wunsch kann wahlweise stationär (am PC) oder mobil (auf ein Aufnahmegerät) diktiert werden. Bei Letzterem wird über ein USB-Kabel das Diktat vom Aufnahmegerät auf den PC überspielt. Die Diktatbearbeitung ist in installierten Standardtextbearbeitungsprogrammen, wie z.B. Microsoft Word, möglich. Somit ist auf Seiten der Schreibkraft kaum besondere Schulung erforderlich. Man benötigt allerdings einen Server auf dem die Spracherkennung stattfindet und auf dem die Audiodaten gespeichert werden.

## **2.) Diktiersysteme in der Praxis**

Der Arbeitsaufwand zur Umsetzung eines Diktats ist immer von der Art des Diktierens abhängig: Man unterscheidet zwischen analogem Diktat (Banddiktat) und digitalem Diktat (Chipdiktat). Bei einem analogen Diktat setzt die Schreibkraft das Diktat auf einem PC – früher: einer Schreibmaschine – um. Nur beim digitalen Diktat ist eine Umsetzung durch den PC mit Hilfe von Spracherkennung möglich. Die Schreibkraft muss den Text dann nur noch in einem Textverarbeitungsprogramm Korrektur lesen. Herr Prahl berichtete, welche Erfahrungen man am Landgericht Detmold und am Amtsgericht Lemgo mit Diktiersystemen gesammelt hat:

Aus organisatorischer Sicht stellt sich zunächst die Frage, wie das Diktat vom Autor zur Schreibkraft gelangen soll. Beim Banddiktat wandert das Diktat auf Mikrokassette mit der Akte vom Autor (am Gericht i.d.R. ein Richter) zur Schreibkraft. Anders beim digitalen Diktat: Hier wird die Audiodatei auf dem Netzwerkserver gespeichert und kann dort von der Schreibkraft abgerufen werden, während die Akte im herkömmlichen Umlauf bleibt. Dabei gibt es verschiedene Möglichkeiten die Audiofiles im Netzwerk zu organisieren: Einmal besteht die Möglichkeit das Diktat in eine definierte Ordnerstruktur hochzuladen. Die Zuordnung zur Akte erfolgt dann durch elektronische Begleitinformationen, Anpassung des Filenamens oder einem Vermerk auf dem Aktendeckel. Ein anderes, zukunftsweisenderes Organisationsmodell bindet die Audiofiles an die zugehörige elektronische Akte (Aktenreferenz), d.h. es erfolgt entweder eine manuelle oder eine automatische Zuordnung zur Akte durch Barcode Scanner, optische Zeichenerkennung oder Transponder (RFID). Herr Prahl demonstrierte die Zuordnung einer Audiodatei zu einer Akte mit Transponder. Hardwarevoraussetzungen für dieses Organisationsmodell sind ein Lesegerät (Kosten: 300-400 €) und ein Transponder (Stückkosten: 1 €). Der Transponder wird an der Akte befestigt, kann aber wiederverwendet werden, wenn die Akte ins Archiv zurückwandert.

## **3.) Fragen und Diskussion**

### **a) Diktate während einer Gerichtsverhandlung**

Digitale Diktate während einer Gerichtsverhandlung sind möglich und erzielen grundsätzlich gute Spracherkennungsergebnisse. Probleme ergeben sich allerdings bei Namen von Anwälten und Zeugen. Die Namen von Anwälten die häufig am Gericht auftreten, können in den Wortschatz aufgenommen werden. Bei vielen Zeugen empfiehlt es sich eine Liste aller Namen anzufertigen und diese mit den Kürzeln „Zeuge 1“, „Zeuge 2“ usw. durchnummerieren. Wenn man beim Diktat die festgelegten Kürzel verwendet, kann nach erfolgter Spracherkennung mit dem Textverarbeitungs-Tool „Suchen & Ersetzen“ das Kürzel bequem durch den wirklichen Namen ersetzt werden.

Diktatfiles, die während Gerichtsverhandlungen entstanden sind, müssen – wie auch Mikrokassetten – 6 Wochen aufbewahrt werden.

#### **b) Zusammenarbeit Schreibkraft - Autor**

Die Erfahrung hat gezeigt, dass Schreibkräfte (nach einer Gewöhnungsphase) dem Einsatz von Spracherkennungssystemen aufgeschlossen gegenüber stehen, wenn die Erkennungsraten sehr gut sind. Sind die Erkennungsraten jedoch schlecht, so bevorzugen sie regelmäßig den Text selbst – ohne Einsatz von Spracherkennung – zu schreiben. Daher muss um der Akzeptanz und der Effektivität willen beständig an der Erkennungsrate gearbeitet werden: Um möglichst gute Ergebnisse zu erzielen, sollten bestimmte Wörter vom Arbeitsplatz aus trainiert werden. Wenn die Schreibkraft den Text korrigiert, kann eine solche Weiterentwicklung nur funktionieren, wenn die Schreibkraft den Autor über die Erkennungsrate informiert und dieser sein Profil verfeinert (d.h. Anpassung des Systems an Lautstärke und Typus der Stimme des Autors). Es wird also künftig mehr Kommunikation zwischen Schreibkraft und Autor notwendig sein. Zusätzlich empfiehlt sich der Einsatz von digitalen Diktiersystemen die erlauben, dass das Profil am mobilen Gerät (mit satzweiser Erkennung) erstellt wird. Dies gibt dem Autor unmittelbar Feedback, sodass er seine Diktierweise anpassen kann.

#### **4. ) Ergebnis:**

Als besonders zukunftsweisend haben sich digitale Diktiersysteme erwiesen. Überzeugt haben insbesondere die Komfortabilität (Indexerstellung, beliebiges Einfügen oder Teillösungen, kein lästiges Spulen, Voice-Activation, hohe Tonqualität, Spracherkennung) und die Anpassungsfähigkeit in bestehenden Organisationsstrukturen (Einsatz in Netzwerken, kein besonderes Abspielgerät bei netzwerkintegriertem Schreib- und Korrekturarbeitsplatz erforderlich). Durch die Möglichkeit Diktatfiles auf einem Fileserver zu speichern, ist ein weiterer Schritt in Richtung elektronischer Akte getan.