

## **Arbeitskreis „Digitales Diktieren“**

Ort und Zeit: Freitag, 23.09.2005, 11.00 Uhr, HS 117

Referenten: 1. Herr Richter am Amtsgericht Peter Uhde (AG Reutlingen)  
2. Herr Richter am Verwaltungsgericht Dietmar Kirschbaum (VG Weimar, derzeit als Leiter der Stabsstelle „IT-FachJustiz“ abgeordnet zum Thüringer Obergericht)

Moderation: Frau Rechtsanwältin Daniela Freiheit

Nach einer kurzen Einführung in das Thema durch die Moderatorin des Arbeitskreises, Frau Rechtsanwältin Daniela Freiheit, berichtete zunächst Herr Richter am Amtsgericht Peter Uhde (AG Reutlingen) über seine Erfahrungen im Zusammenhang mit der Einführung PC-gestützter Online-Spracherkennungssysteme (der diktierte Text wird durch den Computer sofort bzw. mit geringer Verzögerung erkannt und am Bildschirm in geschriebenen Text umgesetzt) in den richterlichen Arbeitsalltag beim Amtsgericht Reutlingen. Das Ergebnis vorwegnehmend zeigte sich Herr Uhde gleich zu Beginn seines Vortrages vom Nutzen dieser Technik und der Möglichkeit ihrer erfolgreichen Einführung – auch gegen die üblichen Vorbehalte und Widerstände bei den zu überzeugenden Richtern – überzeugt, betonte aber zugleich, dass es nicht mit der Anschaffung von Soft- und Hardware sowie anschließendem Fünf-Minuten-Training getan sei. Vielmehr komme es darauf an, die Diktierenden von Anfang an mit umfassenden Schulungen zu begleiten und nicht mit der Technik allein zu lassen. Sonst vorprogrammierte Misserfolge würden sonst schnell zu Frust und Ablehnung führen.

Beim Amtsgericht Reutlingen habe man deshalb auch darauf gesetzt, die Neugier der Richter zu wecken, indem jedes Büro mit einem neuwertigen Computer samt Spracherkennungssystem ausgestattet wurde und die Kollegen zu unverbindlichen Schulungen eingeladen wurden. So konnte erreicht werden, dass dort heute von 15 Richtern zwölf die elektronische Spracherkennung nutzen, während in Tübingen aufgrund organisatorisch ungünstigerer Umstände bei ihrer Einführung dieser Prozentsatz wesentlich geringer ausfällt. Daher könne man derzeit beispielsweise bei einer zeitweiligen Überlastung der Schreibkräfte in Stuttgart, Tübingen oder Sindelfingen bereits einen gewissen Rationalisierungseffekt durch eine digitale Verteilung der verbleibenden Diktate an weniger ausgelastete Serviceeinheiten erzielen.

Als Vorteil der sog. Online-Spracherkennung unterstrich Herr Uhde, dass die hierbei in das Computersystem diktierten, von der Software sofort erkannten und am Bildschirm wiedergegebenen Texte durch den zusätzlichen visuellen Eindruck des Diktierten beim Diktierenden sowie der diesem gegebenen Möglichkeit, jederzeit korrigierend einzugreifen, insbesondere an inhaltlicher Qualität gewinnen. Daneben seien aber auch die durch das Wegfallen der sonst nötigen Überarbeitungsschritte zwischen Richter und Serviceeinheit erreichte Zeitersparnis sowie Effizienzsteigerung nicht zu unterschätzen.

Da Voraussetzung für zufrieden stellende Erkennungsraten ein verändertes Diktierverhalten sei, stünden den zuvor genannten Vorzügen zwar vergleichsweise aufwändige Schulungen – im Falle eines Updates auf eine neue Software-Version u.U. auch Nachschulungen – gegenüber. Denn um sich mit der Online-Spracherkennung eine umfassende Nachbearbeitung des erstellten Dokuments tatsächlich ersparen zu können, kommt es in besonderem Maße darauf an, dass der Diktierende den aktuellen Sprachbefehlssatz, auf den die benutzte Software „hört“, auch selbst beherrscht, da anderenfalls durch „Missverständnisse“ des Programms ständig unerwünschte Effekte hinsichtlich des erkannten Textinhaltes oder seiner Gestaltung auftreten könnten und andererseits auch beabsichtigte Formatierungen von Hand eingebracht werden müssten. Anhand der Formatierungsanforderungen einiger typischer

Textbausteine aus der richterlichen Praxis, die in kurzen Live-Diktaten während der Veranstaltung beispielhaft in eine Online-Spracherkennungssoftware eingesprochen wurden, stellte Herr Uhde die dabei auftretenden Schwierigkeiten sehr anschaulich dar.

Angesichts der Tatsache aber, dass Juristen sich ohnehin ständig mit neuen Materien auseinandersetzen müssten, sei das vermeintliche Hemmnis des hierfür nötigen Schulungsaufwandes letztlich nur eine Frage der Motivation. Bislang würden solche Schulungen in Deutschland durch die Händler der Spracherkennungssoftware, die mit deren Aufbau und Funktionsweise selbstverständlich bestens vertraut seien, selbst organisiert und angeboten. Allerdings würden diese sich wiederum selten mit den spezifischen Problemen des Diktierens selbst auskennen, weshalb Herr Uhde vorschlug, sich hierfür um Multiplikatoren aus den Reihen der Richterschaft selbst zu bemühen, die aus eigener praktischer Erfahrung wüssten, welche Probleme beim Umgang mit diesen Spracherkennungssystemen auftreten könnten. An den EDV-Gerichtstag Verein gewandt, sei es nach seiner Meinung möglicherweise auch interessant und überlegenswert, ob man derartige Multiplikatorenschulungen nicht auch im Rahmen der in Merzig zu gründenden Europäischen EDV-Akademie anbieten wolle.

Ist der einzelne Diktierende nach einer gewissen Trainingsphase schließlich selbst mit den besonderen Voraussetzungen eines für die elektronische Spracherkennung brauchbaren Diktates vertraut, könne für die laufende Praxis auch zu einer Offline-Spracherkennung übergegangen werden. D.h., der Diktierende zeichnet seine gesprochenen Texte lediglich noch digital auf, während eine Servicekraft die entstandenen Sounddateien in einen Computer einlesen sowie durch die Software erkennen lässt und anschließend die wenigen noch verbliebenen Fehler selbst korrigiert. Ein Nachteil dieser Methode liege zwar darin, dass hierbei der Diktierende selbst von einer etwa sinkenden Erkennungsrate nicht mehr unmittelbar durch das Programm erfährt, sondern es hierzu eines ständigen kommunikativen Austausches mit der Serviceeinheit bedarf. Jedoch lassen sich beispielsweise Erkennungsprobleme durch eine belegte Stimme bei einer Erkältung sehr einfach durch eine Feineinstellung des Systems beheben.

Als Systemvoraussetzungen für die beim Amtsgericht Tübingen eingesetzte Spracherkennungssoftware Dragon Naturally Speaking (Professional/ Legal) in der Version 8 nannte Herr Uhde ein Gigabyte (Minimum: 512 MB) Arbeitsspeicher, ein möglichst rauscharmes sowie gegen Nebengeräusche unempfindliches Mikrofon (Preis: ca. 80 bis 120 EUR – die der Software beiliegenden preiswerteren Mikrofone seien insoweit oft weniger geeignet). Idealerweise solle man zudem auch eine am USB-Port angeschlossene externe Soundkarte verwenden. Zwar funktioniere oft auch der in das Computersystem schon eingebaute preiswerte Soundchip sehr gut, jedoch sei diese Ausstattung insbesondere bei Notebooks wegen der durch die dort sehr enge Bauweise oft besonders ungünstigen Rauscheigenschaften nur unter Vorbehalt zu empfehlen; allerdings halte sich dieses Problem wiederum bei stromsparender Systemausstattung (z.B. Intel Centrino Prozessor) u.U. in Grenzen.

Ein insgesamt eher ganzheitliches Konzept in punkto Medienbruchfreiheit bei Fachsystemen, Dokumentenmanagement und Vorgangsbearbeitung, mit dem zugleich fachgerichtsübergreifend eine Geschäftsprozess- und Aufwandsoptimierung erreicht werden soll, wird im Freistaat Thüringen verfolgt. Über die dort in den letzten Jahren vorgenommene einheitliche Erneuerung der technischen Ausstattung, bei der besonderer Wert auf eine Effizienzsteigerung in den Bereichen Fachaufgabenerledigung, PC-Betreuung und Justizverwaltung gelegt wurde, referierte Herr Richter am Verwaltungsgericht Dietmar Kirschbaum (VG Weimar), welcher derzeit als Leiter der mit diesen Aufgaben betrauten Stabsstelle „IT-FachJustiz“ zum Thüringer Obergericht abgeordnet ist.

Die informationstechnischen und finanziellen Voraussetzungen für eine solche landesweit homogene Neuausstattung der Fachgerichte seien erst durch den generellen Umstieg im IT-Beschaffungswesen von Kauf auf Leasing geschaffen worden. Dabei habe man von Anfang an das Hauptaugenmerk auf Zukunftsfähigkeit, Systemintegration, Modularität, hohe Funktionalität und Flexibilität gelegt, um generelle Zielvorstellungen wie die Verwirklichung der elektronischen Akte bzw. des elektronischen Rechtsverkehrs im Sinne einer medienbruchfreien digitalen Vorgangsbearbeitung auf innovative und zugleich wirtschaftliche Weise voranzutreiben. Obwohl in Thüringen seit Jahren mit elektronischer Textverarbeitung gearbeitet werde und die Justiz auch bereits seit 1995 vollständig elektronisch vernetzt sei, sollte die bereits vorhandene analoge Diktiertechnik in diesem ganzheitlichen System noch genauso Verwendung finden, wie die papiergebundene Akte, so dass Diktate auch weiterhin auf herkömmliche Weise auf einem an den Aktendeckel gehefteten – analogen oder digitalen – Tonträger in die Serviceeinheit gelangen könnten.

Bei der Frage nach dem für diese Anforderungen richtigen System seien zunächst die grundsätzlichen konzeptionellen Unterschiede zwischen den vom Ansatz her PC-orientierten Einzelplatzsystemen und den Server-Client-orientierten strukturellen Mehrplatzsystemen zu beachten gewesen. Während erstere in ihren historischen Anfängen zwar bereits ein breites Anwendungsspektrum geboten, jedoch noch Probleme bei der Diktiergeschwindigkeit aufgewiesen hätten, seien die letzteren mit ihrem professionellen Anwendungsfokus (z.B. Krankenhäuser) zunächst noch ohne die sofortige Online-Visualisierung bei der Spracherkennung ausgekommen. Inzwischen hätten sich aber beide Systeme aufeinander zu entwickelt, ohne dabei ihre architektonisch-konzeptionellen Unterschiede zu verlieren.

Im Hinblick auf die Akzeptanz bei den Richtern und Servicekräften sowie auf die Geringhaltung von Ersatzinvestitionen sei zudem Wert darauf gelegt worden, dass sich die neu anzuschaffende digitale Diktiertechnik harmonisch und sukzessive in das bestehende System einfügen würde, weshalb der Umstieg so schonend und flexibel wie möglich habe erfolgen sollen (z.B.: digitale Diktiergeräte mit gewohntem Schiebeschalter/ Schreibservice-Ausstattung wie bisher mit Kopfhörer und Fußschalter). Um einen möglichst effizienten Einsatz der Technik zu erreichen, sollten sich die Richter – ohne Einschränkung der Möglichkeit, auch direkt an die Erkennungssoftware gerichtete Sprachbefehle zu verwenden – weiterhin ganz auf den Inhalt der von ihnen abzusetzenden Diktate konzentrieren können, ohne sich etwa um notwendige Korrekturen bei der Spracherkennung oder gar die Adaption des Systems kümmern zu müssen (sog. „Profidiktat“), da sich erfahrungsgemäß das Technik-Interesse der allermeisten Richter in engen Grenzen halte.

Durch die getroffene Entscheidung für ein Server-Client-orientiertes strukturelles Mehrplatzsystem unter Verwendung von mobilen und stationären digitalen Diktiergeräten seien nunmehr nach wie vor Diktate ohne elektronische Erkennung möglich; ebenso könne aber auch mit Hintergrund- bzw. Batcherkennung oder auch mit Onlineerkennung gearbeitet werden und zugleich eine zentrale Adaption sowie Administration zur Systempflege erfolgen, wodurch volle Flexibilität erreicht werde. Dabei benötige das System allenfalls eine zweiminütige Grundadaption, um bereits eine beachtliche Erkennungsrate erzielen zu können. Die Entscheidungshoheit für den Umgang mit dennoch schlecht erkannten Diktaten liege aber in jedem Falle bei der Servicekraft; diese müsse solche Texte nicht mühsam korrigieren, wenn sie ihn selbst schneller abhören und in Gänze von Hand neu eingeben könne. Zudem biete sich durch die Integration in das vorhandene Computernetzwerk auch die Möglichkeit einer flexiblen Zuweisung der zu bearbeitenden Diktate bei zeitweiliger Überlastung einzelner Servicekräfte an. Aufgrund der digitalen Vernetzung könnten auch kollektive Wortschatzlexika dezentral durch alle angeschlossenen Anwender hinterlegt und weiterentwickelt werden.

Bis voraussichtlich Ende des Jahres werde noch an der Einbindung des digitalen Diktiersystems in die bestehenden Fach- und Dokumentenmanagementsysteme gearbeitet,

während die vollständige Umstellung auf den elektronischen Rechtsverkehr bis 2008 abgeschlossen sein sollte. Wie Herr Kirschbaum während seines Vortrages mit Unterstützung durch seine Mitarbeiterin Frau Rupert demonstrierte, ermöglicht es die verwendete Technik aber beispielsweise schon jetzt, den digitalen Diktiergeräten das jeweilige Aktenzeichen per Sprachkommando als Metadatum mitzugeben (ein Zusatzkürzel teilt dabei dem Erkennungssystem sogar den entsprechenden Dokumententyp mit, wodurch dieses künftig z.B. auch automatisch die für ein Urteil oder einen Beschluss richtige Rubrums-Vorlage auswählen könne) sowie eine zentrale Vorkonfiguration aller digitalen Diktiergeräte auf ihren jeweiligen Autor und die diesem zugewiesenen Serviceeinheit(en) vorzunehmen. Somit könne sich jeder Richter ganz auf sein „Profidiktat“ konzentrieren, diesem anschließend je nach Bedarf eine entsprechende Priorität zuweisen und es durch einfaches Andocken von jedem Arbeitsplatz aus sofort in den Jobpool der richtigen Serviceeinheit übermitteln, wo es über ein Dokumentenmanagementsystem automatisch der richtigen elektronischen Akte zugewiesen und nach Dringlichkeiten abgearbeitet, aber auch in seinem Bearbeitungsstatus weiterhin durch den Richter verfolgt werden könne.

Der im Anschluss an beide Vorträge durch einen Zuhörer geäußerten Besorgnis, dass bei der Offlineerkennung eines Profidiktates der Arbeitsaufwand lediglich vom Diktierenden auf die Schreibkraft verlagert werden könnte, wenn diese den erkannten Text dennoch aufwändig nachbearbeiten oder im Zweifel gar vollständig neu abtippen müsse, hielt Herr Kirschbaum entgegen, dass bei dem Thüringer Projekt der Fokus zunächst weniger auf der digitalen Spracherkennung als viel mehr auf einer schonenden Integration der zukunftsfähigeren digitalen Diktiertechnik in die bestehenden Systeme lag. Zudem ließe sich durch Schulungen, welche sich aber nach seiner Meinung vor allem mit den nicht technisch orientierten Fragen des Diktierens selbst befassen sollten, auch das Potential bei der Spracherkennungsrate weiter steigern.

Dem fügte Herr Uhde hinzu, dass die Fehlerquote bei Online-Spracherkennungssystemen heute im Schnitt bei etwa 3 bis 5 Prozent liege. Insbesondere müsse aber nach ihrer Einführung auch im Falle einer Präadaption mit einer ein bis zwei Wochen andauernden Frustrationsphase gerechnet werden, in welcher nicht jedes miserabel erkannte Diktat sofort verworfen werden dürfe, da das System in dieser Zeit erst durch die manuellen Korrekturen des Bearbeiters noch „lerne“, die in der Aussprache gerade des jeweiligen Diktanden liegenden Besonderheiten zu verstehen.

Abschließend dankte Frau Freiheit beiden Referenten für ihre Ausführungen und wünschte ihnen weiterhin viel Erfolg bei der Fortführung ihrer Projekte sowie der noch zu gründenden Europäischen EDV-Akademie die Realisierung der durch Herrn Uhde vorgeschlagenen Schulungen.

Enrico Krüger