

G. Geidel, G. Saeltzer

# Ist automatische Spracherkennung jetzt praxistauglich oder nicht?

Klinik am Tharandter Wald  
Niederschöna

## Zusammenfassung

### Grundlagen, Testergebnisse, Erfahrungen und Tipps für den Einsatz:

Schlüssel jeder erfolgreichen Kooperation in der Patientenversorgung ist die leistungsfähige Kommunikation der Partner. Kosten- und Termindruck, steigende Qualitätsforderungen sowie Wettbewerb zwingen zu Neuerungen wie dem Einsatz von Spracherkennungssoftware. Diese verwandelt den bisher „taubstummen“ Personalcomputer (PC) in einen hörenden und sprechenden Schreibautomaten. Sein Gebrauch ist natürlich und einfach, aber ungewohnt

und oft überraschend: Sie diktieren in ein PC-Mikrofon und der PC schreibt sofort auf den Bildschirm, was er „verstanden“ hat. Ihre Hände bleiben frei, Sie schonen Finger und Augen und sparen vielleicht Zeit und Geld.

Ist diese Neuerung jetzt für die Arzt- und Klinikpraxis tauglich oder nicht? Ein Praxistest durch in Spracherkennung ungeübte Klinikärzte und den Geschäftsführer einer Rehabilitationsklinik sollte das beantworten. Der Beitrag erläutert für das Verständnis unerlässliche Grundlagen, stellt die inter-

essanten Ergebnisse dar, gibt einige Hinweise und Tipps für den Einsatz. Erfolgreich können Ärzte Spracherkennungssoftware nutzen, sofern sie bestimmte objektive und subjektive Voraussetzungen erfüllen und eine Art „Symbiose“ mit dem eigenwilligen Schreibautomaten eingehen.

**Schlüsselwörter:** Schreiben von Gutachten, Arztbriefen, Befunden, Automatische Spracherkennung, Linguatex Voice Pro 10 Medical Edition, Praxistest durch Ärzte

## Was ist automatische Spracherkennung (ASE) und was leistet sie?

Die Anwendung von ASE am PC unterscheidet sich grundlegend vom Gebrauch herkömmlicher Arzt-/Kliniksoftware. ASE ist automatische, intelligente biometrische Verarbeitung der Stimme des Sprechers. Ein hochwertiges Mikrofon, meist an einem Kopfbügel getragen (Headset), erfasst die Schallwellen – ein Mix von Umgebungs-, Atemgeräuschen und Sprachlauten – und übersetzt diese in lesbaren Text (STT-Systeme, Abbildung 1). Umgekehrt kann der PC mit Computerstimme auch ausgewählten Text vorlesen. Wie mit einem Diktiergerät lässt sich außerdem ein Diktat aufnehmen und abspielen.

Die heutige Leistungsfähigkeit soll ein Gedankenversuch zeigen: Benutzt wird die auf den Nutzer angepasste ASE-Software Dragon NaturallySpeaking 7 preferred (mit dem Qualitätsurteil „gut (2,4)“ nach Stiftung Warentest 1/2004) auf einem Notebook mit USB (Universal Serial Bus)-Mikrofon, das mit einem Flachstecker am PC angeschlossen ist.

Sie setzen am PC ein Headset auf, starten ASE und möchten einen Fachtext in die Schreibsoftware MS Word diktieren (Ärzteblatt Sachsen 8/2004, Seite 368). Sie diktieren (Tabelle 1)

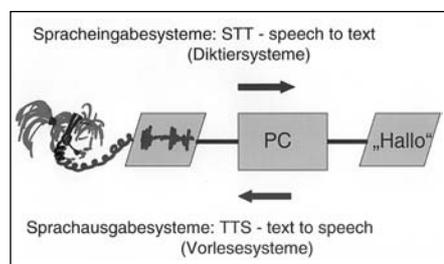


Abb. 1: Sprachein- und Sprachausgabesysteme

mit Satzzeichen (KOMMA) und Absatzangabe (NEUER ABSATZ). Das sind Sprachbefehle, die in das Textdiktat eingefügt sind. Dann holen Sie tief Luft und blicken auf den Bildschirm. Dort steht der erkannte Text (Spalte 2). Beim ersten Vergleich mit dem Diktat stellen Sie fest: alles richtige Wörter (keine ärgerlichen Tippfehler wie „sit“ statt „ist“). ASE schreibt nur richtige Wörter (im Unterschied zur menschlichen Schreibkraft). Bei genauerer Prüfung ist Ihre Frustration und Sinn für Humor gefragt: statt dem diktieren „sein Ruf“ schrieb ASE „Anruf“, statt „rehabilitativ“ steht dort „lernt Digital tief“. No body is perfect! Das gilt auch für den Schreibcomputer. Die ausgeführten Handkorrekturen zeigt Spalte 3.

## Welche Schreibgeschwindigkeit und Erkennungsgenauigkeit lässt sich erreichen?

Dieses Diktat und die Niederschrift durch den PC dauerten etwa eine Minute, die Schreibgeschwindigkeit betrug damit 600 Anschläge pro Minute. Abbildung 2 zeigt Vergleichswerte. Der typische PC-Nutzer hämmert je Sekunde im „Adler-Such-System“ etwa 2 Anschläge auf die Tastatur, eine versierte Schreibkraft 4 Anschläge. Adenauer sprach mit einem Output von 800 Zeichen pro Minute, Nachrichtensprecher bringen es auf 1100. Eine ASE nach Stand der Technik erreicht Adenauers Kommunikationsgeschwindigkeit, also ein Vielfaches eines durchschnittlichen Tastenschreibers.

Die erzielte Schreibleistung von ASE muss man in ein richtiges Licht rücken. Nehmen Sie sich etwas Zeit für diesen kleinen Workflow-Versuch:

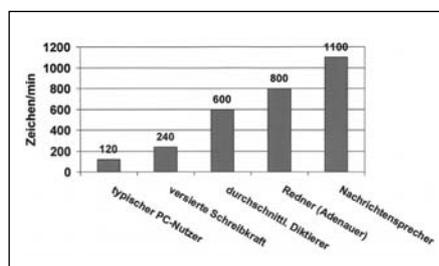


Abb. 2: Kommunikationsgeschwindigkeiten typischer Schreiber und Sprecher

1. Tippen Sie den Text (Tabelle 1) einmal selbst ab und stoppen Sie die dafür benötigte Zeit.
2. Falls Sie Zugriff auf ein Sekretariat haben: Liefern Sie den Text ab zum Schreiben/ Speichern auf einer Diskette. Stoppen Sie die Zeitdauer bis zum Eintreffen des Geschriebenen.
3. Oder stoppen Sie den Durchlauf mit Ihrem Banddiktat (Die Uhr läuft ab Diktatbeginn). Dann vergleichen Sie die gemessene Dauer mit der ASE-Minute. Denken Sie bitte daran, dass das Kostbarste – für den Patienten und Sie selbst unsere Lebenszeit ist. Vergeudete Zeit ist für immer dahin...

Die Genauigkeit des Erkannten lässt sich messen durch das Verhältnis richtiger Tastenschläge (einschließlich Leertaste) zur Gesamtlänge des Textes. Nach Tabelle 1 betrug sie 95 Prozent, die Zeichenfehlerrate 5 Prozent (auch der menschliche Schreiber vertippt sich in dieser Größenordnung). Diese Erkennungsgenauigkeit schaffte das System ohne vorheriges Training solcher Fachbegriffe wie „rehabilitativ“. Nach einem kurzen Zusatztraining erkannte das System sofort

99,5 Prozent des Diktats korrekt, es fügte nur an unpassender Stelle die vier falschen Zeichen „Das“ ein – seine „Übersetzung“ eines Atemgeräuschs des Sprechers.

### Was steckt hinter dem überraschenden Verhalten von ASE?

Die ASE verhält sich immer wieder überraschend, eine Folge ihrer intelligenten biometrischen Messwertverarbeitung und der Eigenart der menschlichen Stimme. Diese ist einzigartig und identifiziert den Sprecher, ist aber abhängig von Stimmung, Gesundheitszustand unter anderen Abbildung 3 zeigt die Schallpegelmessung einer Sprechprobe. Diktieren wurden zweimal „sächsische Gerichte“, dazwischen lag eine Sprechpause. Das Diagramm der 2. Sprechprobe unterscheidet sich signifikant von der ersten: dasselbe wurde gesagt – jedoch mit etwas mehr Energie/Nachdruck (das wird gut sichtbar). Gefühle/Stress verändern unbewusst und sofort den Klang unserer Stimme – ebenso langes Sprechen, Heiserkeit oder Rauchen. Das verkompliziert den Gebrauch und beeinflusst die Erkennungsgenauigkeit. Eine klanglich

veränderte Stimme ist für die ASE eine andere Stimme; sie versteht weniger richtig und mehr falsch.

Der Erkennungsvorgang im PC basiert auf einem Vergleich der registrierten Klänge (Abbildung 3) mit einer Datenbank von Referenzmustern für Klänge/Wortteile/Wörter. ASE erkennt nur Wortklänge und entsprechende Wörter, die der Hersteller im Vokabular mitgeliefert und/oder der Benutzer der ASE in einem speziellen Training (durch Vorlesen vorgegebener Texte) beigebracht hat. ASE verarbeitet und speichert riesige Datenmengen. Wurde die Sprechprobe von 3 Sekunden Dauer mit 11 KHz digital abgetastet, so sind das 33.000 Messwerte, etwa 66 KByte. ASE verdichtet diese in 40 sinnvolle Zeichen, fast auf Zweitausendstel. ASE ist sensibel und störungsempfindlich. Ihre Beherrschung erfordert praktische Anleitung, Erfahrung und „Fingerspitzengefühl“. Ein leistungsfähiger PC ist nötig, ein Multimedia-PC, wie er seit 2000 auf dem Markt ist (Prozessor möglichst schneller als 1 GHz, Arbeitsspeicher mit 256 MB-RAM, qualifiziertes Audiosystem).

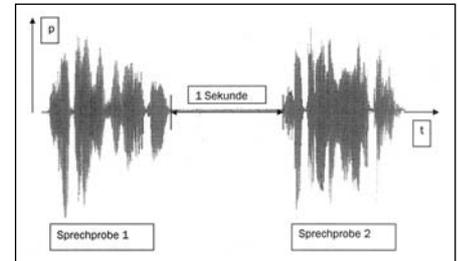


Abb. 3: Mit Mikrofon gemessenes Schallpegel(p)-Zeit (t)-Diagramm zweier Sprechproben mit einer Sprechpause

### Welche Motive bewegen zum Einsatz?

Verschiedene Gründe führen zum Einsatz von ASE:

- Zeit sparen (Arbeits- und Lebenszeit)
  - Kosten senken, Geld sparen
  - Arbeitsabläufe beschleunigen
  - Hände frei bekommen, um Zusätzliches zu erledigen
  - Qualität der Arbeitsvorgänge und ihre Ergebnisse verbessern
- sowie
- schneller und natürlicher kommunizieren

#### 1. Das wurde diktiert

Zitat aus dem Ärzteblatt Sachsen 8 SCHRÄGSTRICH 2004 KOMMA Seite 368 NEUER ABSATZ

Berufspolitik

NEUER ABSATZ

Benötigen wir einen Arzt neuen Typs FRAGEZEICHEN NEUER ABSATZ

Das deutsche Gesundheitswesen ist nach wie vor besser als sein Ruf KOMMA und viele Staaten beneiden uns um dieses System gesundheitlicher Betreuung PUNKT Ohne wesentliche Wartezeiten wird die Bevölkerung wohnortnah und flächendeckend ambulant KOMMA stationär und rehabilitativ versorgt PUNKT Oberste Priorität aus Sicht der Patienten hat GEDANKENSTRICH jüngsten Meinungsumfragen zufolge GEDANKENSTRICH der ungehinderte freie Zugang des Patienten zu seinem Arzt PUNKT Dort haben kranke Menschen Anspruch auf eine fachlich hochwertige und humane Versorgung PUNKT

#### 2. ASE verstand und schrieb

Zitat aus dem Ärzteblatt Sachsen 8/2004, Seite 368

Berufspolitik

Benötigen wir einen Arzt neuen Typs?

Das deutsche Gesundheitswesen ist nach wie vor besser als Anruf, und viele Staaten beneiden uns um dieses System gesundheitliche Betreuung. Ohne wesentliche Wartezeiten wird die Bevölkerung Wohnorten war und flächendeckend ambulant, stationär und lernt Digital tief versorgt. Oberste Priorität Aussicht der Patienten hat – jüngsten Meinungsumfragen zufolge – der ungehinderte freie Zugang des Patienten zu seinem Arzt. Dort haben kranke Menschen Anspruch auf eine fachlich hochwertige und humane Versorgung.

#### 3. Korrigierter Text

Zitat aus dem Ärzteblatt Sachsen 8/2004, Seite 368

Berufspolitik

Benötigen wir einen Arzt neuen Typs?

Das deutsche Gesundheitswesen ist nach wie vor besser als **sein** Ruf, und viele Staaten beneiden uns um dieses System gesundheitlicher Betreuung. Ohne wesentliche Wartezeiten wird die Bevölkerung **wohnortnah** und flächendeckend ambulant, stationär und **rehabilitativ** versorgt. Oberste Priorität **aus** Sicht der Patienten hat – jüngsten Meinungsumfragen zufolge – der ungehinderte freie Zugang des Patienten zu seinem Arzt. Dort haben kranke Menschen Anspruch auf eine fachlich hochwertige und humane Versorgung.

Tabelle 1: Demonstrations-Diktat zum Vergleich. Spalte 1: Das Wortdiktat (mit Satzzeichen und Absätzen in GROSSBUCHSTABEN) Spalte 2: Von ASE automatisch erkannt und geschriebener Text Spalte 3: Text mit den per Hand korrigierten Erkennungsfehlern (in Fettschrift)

- Gesundheit schonen, RSI (repetitive strain injury, sog. „Mausarm“) vorbeugen
- Gewinn an persönlicher Freiheit und Flexibilität
- Neue Chancen für Kranke und motorisch Behinderte.

Durch die immer intensivere und unnatürliche Tastenarbeit am PC (auch durch die Handy-Nutzung) und das „Mausjonglieren“ steigt ständig die Anzahl belastungsbedingter Erkrankungen wie RSI. Gegen diese neue Volkskrankheit gibt es nur ein Gegen- und Vorbeugungsmittel: weniger Handarbeit am PC. Also: Diktieren statt Tastentippen und Mausschieben/ Klicken. Der Einsatz von ASE ist eine Maßnahme des vorbeugenden Gesundheitsschutzes, zu dem der Arbeitgeber verpflichtet ist (§ 4 Arbeitsschutzgesetz)!

Wer sich beim Schreiben unabhängig von Schreibdiensten macht (und deren Stimmung und Einsatzbereitschaft), gewinnt dadurch persönlich Freiheit und Zeit. In kritischen Situationen kann die schnellere schriftliche Kommunikation mit ASE Verzögerungen und Missverständnisse bei der Befundübermittlung vermeiden oder wichtige Dokumentation in einem Arzthaftungsprozess liefern. ASE hat Bezug zur Qualitätssicherung und kann auch Gegenstand der Arbeit von Qualitätszirkeln sein. Betriebswirtschaftlich ist der richtige Einsatz von ASE an der richtigen Stelle attraktiv: Er ist eine Investition, die sich am schnellsten rechnet („amortisiert“), statt erst nach Jahren – schon nach Monaten.

#### Was waren das Ziel, die Teilnehmer und die Plattform des Praxistests?

Anstoß für den Test war ein einführender Vortrag über ASE mit praktischer Vorführung in der Rehabilitationsklinik (Obiges ist daraus entnommen). Dort lagen bisher keine praktischen Erfahrungen mit ASE vor – bis auf einen früher gescheiterten Versuch. Die Erfahrungen der Teilnehmer im Umgang mit PC, speziell MS Word, streuten von sehr gering bis viel. Testplattform war die installierte Spracherkennungssoftware Linguattec Voice Pro 10 (Testsieger nach test 1/2004), Version: Medical Edition mit USB-Mikrofon (Komplettpreis etwa 399 €). Zuerst wurde ein Notebook MaxData mit 1,6 GHz Prozessor und 256 MB Arbeitsspeicher und Betriebssystem Windows XP benutzt. Nachdem dieses Notebook bei höheren Zimmer-

temperaturen und längerem Betrieb immer abstürzte, wurde ein handelsüblicher PC neuer Bauart mit Betriebssystem MS Windows 2000 eingesetzt. Angeleitet durch einen erfahrenen ASE-Trainer stellten die 7 Ärzte und der Geschäftsführer einzeln das Audiosystem auf ihre Stimme ein, trainierten das System mit einer Textprobe und führten eine spezielle Vokabularanalyse selbst gewählter Fachtexte aus. Nach diesen Systemanpassungen hatte jeder die Gelegenheit, selbst gewählte Fachtexte wie Arztbriefe, Befunde, Berichte in MS Word zu diktieren und zu korrigieren. Diese Praxistests dauerten jeweils circa 2 Stunden. Die Teilnehmer hielten danach ihre Eindrücke in einem Fragebogen mit 4 Fragen fest:

1. Die Anwendung der Software ist leicht und erlernbar
2. Das System ist für das Schreiben von Arztbriefen geeignet
3. Das System ist für das Diktieren von Befunden geeignet
4. Das System möchte ich für meine Zwecke in der Klinik einsetzen.

Die Teilnehmer konnten folgende Antworten ankreuzen: Ich finde keine Antwort, trifft nicht zu, trifft wenig zu, trifft noch zu, trifft überwiegend zu, trifft voll zu. Die Urteile streuten. Betrachtet werden die Mittelwerte.

#### Wie urteilten die Teilnehmer nach dem Praxistest über die benutzte ASE?

Die Teilnehmer beurteilten die benutzte ASE im Mittel so (Abbildung 4).

1. Die höchste Zustimmung mit 81 Prozent (mit der geringsten Streuung) fand die leichte Erlern- und Anwendbarkeit der Software. Diese Aussage bezieht sich aber nur auf die individuelle Anleitung im Praxistest, belegt deren Zweckmäßigkeit. Das Handbuch sowie die Software-Help-Funktion wurden nicht

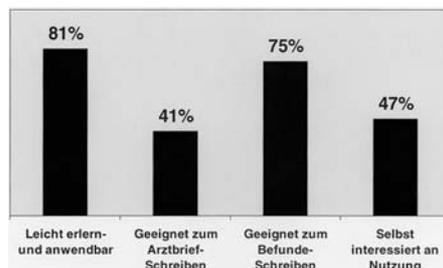


Abb. 4: Mittlere Zustimmung bei der Beurteilung der 4 Kriterien (in Prozent) durch die 8 Teilnehmer des Praxistests von Linguattec Voice Pro 10 Medical Edition

benutzt (sie enthalten leider die üblichen Unklarheiten und Fehler, die ohne Expertenhilfe einen ASE-System-Laien schnell scheitern lassen).

2. Das positive Urteil mit 75 Prozent zur Eignung für Befundeschriften, einem Kernstück der Arztdokumentation, fällt auf – zumal die Probetexte komplizierte medizinische Fachtexte waren. Dem entspräche auch eine positive Zustimmung zu ärztlichen Kurzmitteilungen (Kurzbriefen, gesicherten E-Mails) mit Befunden, aber auch zu Kurzdokumentationen (wurde nicht getestet).

3. Die geringste Zustimmung mit 41 Prozent fand die Eignung zum Schreiben von Arztbriefen. Das steht im Widerspruch zu 2. und lässt sich so interpretieren:

- Es existiert ein funktionierendes EDV-System für das Schreiben von Arztbriefen durch die Sekretariate.
- Die Ärzte möchten sich lieber ihren Patienten zuwenden, als den für Kommunikation und Dokumentation notwendigen, aber lästigen Papierkram am Computer erledigen und schon gar nicht den Job vorhandener Schreibkräfte verrichten (was für die Ärzte und ihren Durchblick spricht).

Die Software bietet Möglichkeiten, längere Standardtexte mittels selbst definierter Kurzbefehle (2 bis 3 Wörter) zu diktieren. Beispielsweise schreibt der Sprachbefehl (in Sächsisch) „Meen Priehfköppel“ den eigenen kompletten Briefkopf mit Logo und E-mail-Adresse (wurde aus Zeitgründen nicht getestet, auch nicht die Qualität der medizinischen Fachvokabulare und die Schnittstelle zu der benutzten Kliniksoftware).

4. Bemerkenswert ist die Zurückhaltung hinsichtlich einer eigenen Nutzung (größte Streuung von allen Kriterien) mit 47 Prozent. Sie wird beeinflusst durch

- sehr unterschiedliche Erfahrung im Umgang mit PC und dem Textsystem MS Word.
- Frust und Ärger über bisherige Software, die den Ärzten oft genug Mehrbelastung statt der versprochenen Entlastung brachte.

Hier kommt die richtige und wichtige Erfahrung der Ärzte zum Ausdruck, dass der professionelle Gebrauch von ASE auch eine gewisse Zeit zum Lernen und Eingewöhnen erfordert (Erfahrungswert: etwa einen Monat lang arbeitstäglich eine 3/4 Stunde Computer-Diktate üben). ASE lässt sich nicht über Nacht beherrschen, ist nicht zum „Nulltarif“

zu haben und damit keine Ausnahme unter Innovationen. Es klingt die Befürchtung an, dass den Ärzten durch die Einführung von ASE weitere Zeit für die Patienten verloren gehen könnte. Dennoch regte der Praxistest einzelne Teilnehmer an, eine private Anschaffung von ASE für ihren Briefverkehr und E-Mails zu prüfen.

Die Teilnahme an den Tests spricht auch für die Aufgeschlossenheit aller Teilnehmer für Innovationen zum Wohl der Patienten und Ärzte.

Wer nicht wie ein Weltmeister rasant in die Tasten zu greifen vermag (oder jederzeit Zugriff auf einen solchen Meister hat), sondern sich selbst langsam, mühsam und ungeschickt auf der Tastatur abplagt, empfindet ASE als Segen, den er nicht mehr im Leben vermissen will.

## **Wie können Sie Fehlinvestitionen und die Vergeudung von Zeit und Geld vermeiden?**

Alle bisherigen ASE-Einsätze zeigten: der erfolgreiche Einsatz ist an diverse technisch-räumlich-organisatorische, aber auch persönliche Voraussetzungen des Nutzers gebunden, wie deutliche und disziplinierte Aussprache beim Diktat, PC-Fertigkeiten und Umgang mit Textsoftware, Kritik- und Frustrtoleranz sowie Kooperationsfähigkeit u. a. (auf interessante Details kann hier aus Platzgründen

nicht eingegangen werden). Der ständige Gebrauch führt zu einem „eheähnlichem Verhältnis“ (Stiftung Warentest) zwischen dem Nutzer und ASE.

Wie kann der Arzt am einfachsten und unkompliziert prüfen, ob ASE für ihn eine echte Alternative ist oder nicht? Wie kann er Fehlinvestitionen an Zeit und Geld vermeiden? Wie kann er selbst seine Eignung für diese Symbiose, die „ASE-Ehe“, prüfen? Versuch macht klug: Machen Sie unter Anleitung eines erfahrenen persönlichen Trainers Ihren persönlichen Praxistest mit typischen Fachtexten Ihres Arbeitsgebietes – wie in dem geschilderten Test. Als Dauer sollten Sie mindestens 2 Stunden vorsehen (für Installation und Deinstallation auf Ihrem eigenen PC zusätzlich mindestens 1 Stunde). So werden Sie ASE als einen Schreibpartner hautnah erleben und können sich eine fundierte Meinung bilden – unabhängig von Hochglanzwerbung oder Schauermärchen. Sie entscheiden selbst, ob ASE zu Ihnen passt, und Sie sich künftig auf eine längere ASE-Ehe einlassen möchten oder nicht.

## **Fazit und Ausblick**

Es hieße Eulen nach Athen tragen, zu erläutern, wie wichtig schnell funktionierende Kommunikation in Notfall- und Risikosituationen und generell beweissichernde Dokumentation ist.

ASE kann durch bessere Kommunikation und Dokumentation die Fähigkeit des Arztes steigern, Leben zu erhalten, Krankheiten zu lindern und zu verhindern oder Gesundheit zu fördern oder einfach im Arzt-Haftungsfall mehr Sicherheit geben. Die entsprechende Aus- und Fortbildung für Ärzte wird zum Muss. Der Einsatz von ASE und die nötige Aus- und Fortbildung sollten aktiv gefördert werden. Auch als wertvolle praktische Starthilfe für junge Ärzte.

Der qualifizierte Einsatz von ASE kann unser Leben – sowohl privat als auch beruflich – signifikant bereichern. Automatische Spracherkennung kann zu einer unersetzlichen Lebens- und Arbeitshilfe werden – sogar mit gesundheitsfördernden Nebenwirkungen.

Autoren:  
Dipl. Ing. Ök. Günter Geidel, Geschäftsführer  
Klinik am Tharandter Wald  
Rehabilitationszentrum Niederschöna GmbH  
Herzogswalder Straße 1  
09600 Niederschöna, OT Hetzdorf

Korrespondenz/Anfragen bitte an:  
Dr. Dr.-Ing. habil. Gerhard Saeltzer  
Postfach 20 01 32  
01191 Dresden  
Telefon: 0162 3592034