

# What's next



Dragon Professional

## Die Geschichte der Spracherkennung – Teil 1

Die Dragon Spracherkennung gibt es bereits seit den frühen 1980er Jahren. Sie war von Beginn an führend und die treibende Kraft, wenn es darum ging, die verfügbare Technologie zur Umwandlung von Sprache in editierbaren Text und die Sprachsteuerung von Computern voranzubringen. Dragon ist gleichermaßen ein Teil der Geschichte und der Zukunft der Spracherkennung. In diesem ersten Teil des zweiteiligen Blogs werfen wir einen Blick auf die frühe Geschichte der Spracherkennung und wie Dragon einen Paradigmenwechsel einläutete. In Teil zwei betrachten wir, wie sich Dragon entwickelt und neue Technologien angenommen, die Genauigkeit erhöht und die Produktivität branchenübergreifend verbessert hat. Bleiben Sie dran!

**Sandra Noetzel**

Posted 9 Juli 2019



Dragon Systems wurde 1982 von Dr. Jim Baker und Dr. Janet Baker gegründet. Sie entwickelten eine Software für Spracherkennung, die gesprochene Worte in Text umwandelte, der daraufhin auf einem Bildschirm erschien. Mit Blick auf die begrenzten Möglichkeiten von Computern jener Zeit war dies eine enorme Leistung. Einige Beispiele zeigen stellvertretend den technologischen Stand von Computern im Jahr 1982:

- Microsoft Windows gab es noch nicht – es kam 1985 auf den Markt.
- Laptops waren selten, groß und teuer – das Modell Grid Compass 1100 kostete etliche tausend Dollar.
- Das Internet existierte zwar, doch das World Wide Web war noch Jahre entfernt. Erst 1990 entwickelte Tim Berners-Lee HTML und machte damit das Erstellen von Websites möglich.

## Die Frühzeit der Spracherkennung

Schon lange vor dem Computerzeitalter gab es Beispiele zur Erfassung der Sprache. So war Alexander Graham Bell, der Erfinder des Telefons, 1881 an der Entwicklung eines Systems beteiligt, das als Reaktion auf Laute einer Stimme Rillen in einen Zylinder aus Wachs ritzte. Im

frühen 20. Jahrhundert folgten Diktiergeräte, deren Aufnahmen erst auf Wachs, dann auf Kunststoff und schließlich mit fortschreitender Technologie auf Magnetbändern aufgezeichnet wurden.

Dies alles diente jedoch nur dazu, Sprache aufzuzeichnen und wiederzugeben. Der große Durchbruch – und das, was wir heute unter Spracherkennung verstehen – kam mit Computersystemen. Hier gab es in den 1950er und 1960er Jahren etliche parallele Entwicklungsstränge. Zum Beispiel:

- 1952 tauchte Bell Labs mit Audrey (Automatic Digit Recognition) auf, das jedoch nur die Zahlen 0 bis 9 erkennen konnte. Der Sprecher musste nach jedem Wort eine Pause machen und Audrey musste auf die Stimme des Sprechers trainiert werden. Aber es funktionierte. Audrey „erkannte“ die Zahl und ließ ein entsprechendes Licht aufleuchten.
- 1962 enthüllte IBM Shoebox auf der Weltausstellung. Shoebox konnte 16 englische Wörter verstehen. Es sollte die Worte hören und einen Befehl vervollständigen, zum Beispiel Zahlen addieren und das Ergebnis liefern.

Die unterschiedlichen Entwicklungen jener Zeit basierten darauf, dass gesprochene Worte zu Stimmenmustern passten. Sie konnten nur Wort für Wort verarbeiten und waren nicht in der Lage, ganze Sätze zu produzieren.

Ein weiterer Durchbruch gelang 1971 mit Harpy. Das Projekt wurde durch die DARPA finanziert (kurz DARPA) (frei übersetzt: Organisation für Forschungsprojekte der Verteidigung – USA) und war ein Gemeinschaftsprojekt der Carnegie Mellon University, des Stanford Research Institute und IBM. Harpy konnte alltägliche Sprache verarbeiten und einzelne Wörter auswählen, hatte jedoch nur einen Wortschatz von rund 1.000 Wörtern.

## Feuer frei mit Dragon

Den größten Fortschritt gab es jedoch 1982, als Dr. Jim Baker und Dr. Janet Baker mit Dragon Systems starteten und den Prototypen eines Spracherkennungssystems vorstellten, das auf mathematischen Modellen basierte. Die Bakers waren Mathematiker und ihr System basierte auf einem Hidden Markov Model (HMM/verborgenes Markow-Modell), das auf Statistiken zurückgreift, um Worte, Sätze und Ausdrücke vorherzusagen.

Damit war weitaus mehr möglich als nur Wörter zu identifizieren. Es erlaubte auch, mit Satzbau und Inhalten zu arbeiten. Dies ist enorm wichtig für eine effiziente, uneingeschränkte Spracherkennung, die in der Lage sein soll, sinnvolle Sätze zu erzeugen. Um beispielsweise einen grammatikalisch korrekten Satz zu erzeugen, muss die Spracherkennung zwischen verschiedenen, gleichklingenden Wörtern und deren unterschiedlicher Bedeutung und/oder Buchstabierung unterscheiden können.

1990 kam mit Dragon Dictate das erste zur allgemeinen Verwendung geeignete Diktiersystem

mit umfangreichem Wortschatz auf den Markt, das Sprache in Text verwandelte. Für Dragon war dies ein bahnbrechendes Produkt, es verlangte von den Anwendern jedoch eine kurze Pause zwischen den einzelnen Wörtern. Dieses Problem war 1997 überwunden. In jenem Jahr kam Dragon Naturally Speaking v1 auf den Markt. Von nun an war kontinuierliche Spracherkennung möglich – Anwender konnten in natürlicher Art und Weise sprechen, ohne Pausen zwischen einzelnen gesprochenen Wörtern machen zu müssen.

In [Teil zwei](#) betrachten wir, wie sich Dragon entwickelt und neue Technologien angenommen, die Genauigkeit erhöht und die Produktivität branchenübergreifend verbessert hat.

**Tags:** [Dragon](#), [KI](#), [Künstliche Intelligenz](#), [produktivität](#), [Spracherkennung](#)



## About Sandra Noetzel

Sandra Noetzel verantwortet das Marketing für Dragon Professional & Consumer (P&C) innerhalb des Geschäftsbereiches Healthcare für DACH und Frankreich. Mit 20 Jahren Erfahrung im B2B- & B2C-Marketing hat es sich Sandra zur Aufgabe gemacht, Spracherkennung als integralen Bestandteil des modernen Arbeitsplatzes in Unternehmen zu etablieren und neue, flexible Arbeitsweisen durch cloud-basierte Softwarelösungen zu ermöglichen. Die ehemalige Tennisspielerin hat an der University of Nebraska, Lincoln und in Oxford, England Werbung/ Integrierte Marketing Kommunikation und internationale Ökonomie studiert. Sie liebt College Football, County Music und die Sonne.

[View all posts by Sandra Noetzel](#)