

Gebührlische Vorlesung

Die Firma Loquendo hat ein Software Development Kit entwickelt, das Texte in verschiedenen Sprachen und Stimmlagen vorliest. Als lizenzkostenpflichtiges Produkt muss es sich mit verwandten freien Projekten vergleichen lassen. Carsten Schnober



Schon lange geben Computer geschriebene Texte per Sprachsynthese wieder, meist jedoch mehr schlecht als recht. Loquendo [<http://www.loquendo.com>], eine Tochterfirma der italienischen Telecom, verkauft ein Software Development Kit, das derartigen Anwendungen zu einer phonetisch besseren Ausgabe verhelfen soll.

Deutschsprachig

Eine Spezialisierung auf die eingesetzte Sprache ist bei dieser Aufgabenstellung unerlässlich. Deshalb liefert Loquendo Stimmen für 15 Sprachen, darunter Englisch und Deutsch und die meisten anderen westeuropäischen Sprachen sowie Chinesisch. Auch Varianten des Englischen und Spanischen aus verschiedenen Teilen der Erde stehen zur Verfügung. Zusätzlich gibt es mehrere Stim-

men zur Auswahl, und zwar in unterschiedlichen Tonfällen, sowohl männliche als auch weibliche.

Für englische Texte gibt es bereits eine relativ breite Programm-Auswahl. Der Schwachpunkt der beiden artverwandten freien Projekte Mbrola [<http://tcts.fpms.ac.be/synthesis/mbrola.html>] sowie Festival [<http://festvox.org/festival>] ist Deutsch. Für Mbrola gibt es zwar deutsche Wörterbücher, sie erfordern aber noch eine Menge Handarbeit, um Texte verständlich zu machen. Auch für Festival gibt es eine deutsche Erweiterung, allerdings keineswegs auf dem aktuellen Stand.

Der Vorteil von Loquendos Text-to-Speech-System liegt in seiner Flexibilität. Während man mit Mbrola trotz Wörterbuch jeden einzelnen Laut phonetisch beschreibt, liest das Loquendo TTS geschriebene Texte ohne weitere Hilfe vor; so verarbeitet es auch maschinell generierten Input weiter. Mit dem Development Kit lassen sich daher auch anspruchsvollere Anwendungen entwickeln, beispielsweise für das automatische Vorlesen von E-Mails.

Das allgemeine API ist unabhängig vom Betriebssystem, zusätzlich finden Windows-Entwickler auch ein Windows-Speech-API. Ein über hundert Seiten starkes Handbuch zum Ausdrucken im PDF-Format dokumentiert die Funktionen des API präzise und mit Hilfe von kurzen Beispielen.

Die bekannten Probleme

Die Loquendo-Software kämpft mit denselben Problemen wie jede andere in der Welt der professionellen Sprachsynthesizer: Zwar ist die Aussprache der einzel-

nen Laute meist korrekt, die Betonung bei verschachtelten Sätzen und zusammengesetzten Wörtern erfordert aber mitunter genaues Hinanhören.

Ein anderes Problem, das solche Software naturgemäß gerade im Bereich elektronischer Kommunikation hat (ob E-Mail oder SMS), bleibt gänzlich ungelöst: Abkürzungen, Fremdwörter beispielsweise aus dem Englischen und ähnliche Spezialfälle ergeben einen unverständlichen Wirrwarr.

Zum Probehören haben die Tester zwei durchschnittlich schwierige Sätze aus der aktuellen Ausgabe des Linux-Magazins synthetisiert; sie stehen unter [<http://www.linux-magazin.de/Service/Listings/2004/11/loquendo>] für den Download zur eigenen Meinungsbildung zur Verfügung.

Kauf

In Deutschland übernimmt die Firma Inloq [<http://www.inloq.de>] den Vertrieb des Loquendo TTS. Die Kosten gestalten sich individuell, und zwar je nach Art der Anwendung, die mit dem SDK entwickelt wird, und liegen im Bereich bis zu 1000 Euro pro Lizenz, Mengenrabatte sind möglich.

Der Preis scheint zunächst für eine bei weitem nicht perfekte Programm-Bibliothek überhöht. Vergleicht man die eingeschränkten Fähigkeiten jedoch mit denen ähnlicher Software, wird deutlich, dass dieses Produkt unter Linux derzeit konkurrenzlos ist und dem aktuellen Stand der Sprachsynthese entspricht. Wer also eine Software für die typischen Einsatzorte wie Mailboxen entwickeln möchte, könnte mit dieser Investition durchaus weiterkommen. ■