

NewtonDictate: Přepis řeči v češtině!

Díky aplikaci NewtonDictate můžete svou skutečnou zapisovatelku nahradit elektronickou verzí. **POČÍTAČOVÝ PŘEPIS ŘEČI** nyní funguje i v češtině.

MICHAL ČERNÝ

Oblast strojového zpracování lidské řeči má i v současnosti stále příchutí něčeho futuristického, náležitějšího spíše do oblasti vědecké fikce. Již dnes ale existuje několik komerčních produktů, které tuto funkci nabízejí – můžeme si vzpomenout na společnost IBM, oddělení Voice Technologies and Systems, na vedoucího tohoto oddělení Jana Šedivého a produkt ViaVoice, který se používá například pro rozpoznání řečových příkazů v automobilech. Tyto produkty mají pro českého uživatele jednu nevýhodu – většinou nejsou určeny pro rozpoznávání řeči v češtině. Problematika rozpoznávání češtiny je mimořádně komplikovaná – zatímco v angličtině si nástroje vystačí se slovníkem čítajícím kolem 50 000 slov, čeština je flexivní jazyk (jedno slovo může mít mnoho tvarů – například slovo „důležitý“ kolem čtyřiceti), takže velikost slovníku může činit až jeden milion slov. Z toho důvodu ne-

lze snadno převzít řešení z jiných jazyků, za nimiž stojí firmy Philips, IBM, Microsoft, Nuance a další.

Vývoj produktu NewtonDictate trval přibližně patnáct let, jeho autorem je Technická univerzita v Liberci, přesněji oddělení SpeechLab pod vedením profesora Jana Nouzy. Poslední tři roky pak na vývoji spolupracovala firma Newton Technologies (dříve Newton Media).

Rozpoznání běžné řeči: Téměř nepoužitelné

Tolik oficiální vyjádření – a nyní se dostáváme k podstatné otázce: Do jaké míry je tato elektronická zapisovatelka skutečně využitelná v praxi? Pokud byste si NewtonDictate nainstalovali, zjistili byste, jak je pro onu značně hloupou, neinteligentní škatuli, jakou je počítač, úkol „speech recognition“ náročný. Kvalita rozpoznání závisí na volbě slovníku (na výběr je všeobecný, právnícký či medicínský slo-

NEWTON DICTATE 1.5

Nástroj pro přepis mluvené řeči do textu

- Do obecného slovníku je zahrnuto 330 000 slov (zajišťuje 98% pokrytí textů).
- Pravidelná aktualizace slovníku (výrazy „Obama“, „Đuričko“, „heparinový“).
- Právnícký oborový slovník zahrnuje 300 000 slov, lékařský 100 000 slov.

Hardware: Program funguje na běžně dostupných počítačích, které by neměly být starší než dva roky. Kromě mikrofonu není nutné žádné speciální vybavení.

Cena: 24 600 Kč


Info: www.diktovani.cz

vník) a hlavně na celkové srozumitelnosti hlasového záznamu. Nástroj jsme vyzkoušeli na záznamu běžné řeči (tedy značně ledabylé, s mnoha „zašumlovanými“, nesrozumitelnými výrazy), a musíme konstatovat, že pro rozpoznání řeči, jakou se mluví v běžném životě, je program takřka nepoužitelný. Jiná situace nastává, mluví-li se do mikrofonu skutečně zřetelně a srozumitelně, používají-li se specializované slovníky a je-li program na váš hlas adaptovaný (součástí NewtonDictate je také volba, která nástroj umožňuje přizpůsobit vašemu hlasu).

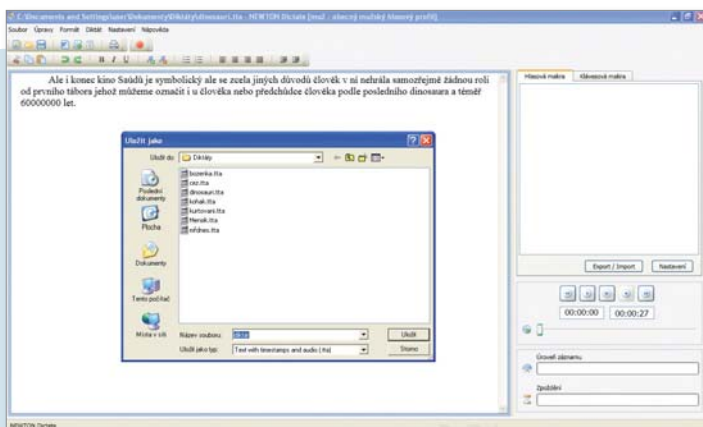
Dinosauri, nebo Saúdové?

Za použití všeobecného slovníku a bez předchozího tréninku program rozpozná text většinou s několika nepřesnostmi. Původní úryvek z knihy zněl takto: „I konec dinosaurů je symbolický, ale ze zcela jiných důvodů. Člověk v něm nehrál samozřejmě žádnou roli. Od prvního tvo- ra, jehož můžeme označit jako člověka nebo předchůdce člově-

ka, odděluje posledního dinosaura téměř 60 milionů let.“ Z NewtonDictate pak vypadl text ve znění: „I konec **kino Saúdů** je symbolický ale se zcela jiných důvodů. Člověk v ní nehrála samozřejmě žádnou roli. Od prvního **tábora** jehož můžeme označit i u člověka nebo předchůdce člověka podle posledního dinosaura a téměř 60000000 let.“

Je tedy zřejmé, že pro diktování je nástroj použitelný s menšími následnými úpravami, lepších výsledků pak lze dosáhnout pomocí hlasové adaptace a s použitím specializovaných slovníků. A jaká je cena a dostupnost softwaru? NewtonDictate je v prodeji už od listopadu loňského roku, základní verze včetně roční aktualizace stojí 24 600 korun. „Používají jej zejména soudci, advokátní kanceláře, lékaři, překladatelé, zapisovatelky. Jednáme také s policií, která by mohla program dobře využít,“ uvádí Petr Herian ze společnosti Newton Technologies. 

MICHAL.CERNY@CHIP.CZ



Synchronizace zvuku a textu: Aplikace ukládá rozpoznávaný text ve speciálním formátu (TTA), který uchovává propojení mezi psaným textem a mluvenou řečí.