

Pracujeme na slovníku s univerzální slovní zásobou říká manažer firmy, která stojí za vývojem systému pro rozpoznání řeči NovaVoice

David Povolný

Společnost Consulting Company Novasoft nabízí ve spolupráci se Západočeskou univerzitou v Plzni systém automatického přepisu mluvené řeči do počítače – program NovaVoice. V rámci projektu eJustice představila v květnu 2008 soudní slovník na převod řeči do textu, který byl vyzkoušený na vybraných soudech v Česku. V říjnu 2008 také obdržela ocenění Křišťálový disk v kategorii profesionální software na veletrhu Inxev, v listopadu 2009 pak získala ocenění Česká hlava. Podle předsedy představenstva společnosti Vladimíra Moravce se software používá hlavně ve zdravotnictví a v justici.

Jaká je historie vašeho produktu?

Na programu NovaVoice spolupracujeme se Západočeskou univerzitou, katedrou kybernetiky na fakultě aplikovaných věd, jmenovitě s profesorem Psutkou a docentem Müllerem. Naše spolupráce začala v roce 2007, nicméně náš produkt vychází z více než dvacetiletého tamějšího výzkumu, který má vysoké renomé. Už v roce 2001 byl výzkumný tým přizván k řešení velmi prestižního projektu MALACH, finančně podporovaném americkou grantovou agenturou. Cílem projektu byl automatický převod rozsáhlého archivu výpovědí svědků holocaustu do textu, a to zejména pro účely rychlého vyhledávání. Vědci v průběhu šestiletého projektu vyvinuli systémy pro převod do textu z mluvené češtiny, ruštiny, slovenštiny a polštiny. Kvalitu jejich výzkumu v oblasti rozpoznávání řeči dokumentuje i to, že už rok zajišťují pro Českou televizi pilotní projekt titulkování živých přenosů z jednání Poslanecké sněmovny Parlamentu ČR.

V jakých oblastech nachází váš software nejčastěji uplatnění?

NovaVoice jsme od počátku koncipovali jako profesionální diktovací nástroj pro oblasti, kde se diktují originální texty. Favority mezi obory byly zdravotnictví a justice. Postupně jsme vytvořili několik slovníků pro řadu zdravotnických oborů. Lékaři používají náš software na diktování lékařských zpráv a nálezů, protože zrychluje jejich práci a zkvalitňuje výsledný přepis. V oblasti soudnictví jsme ve spolupráci s Krajským soudem

v Praze vytvořili slovník pro diktát soudních rozhodnutí v usneseních a rozsudcích. Na základě poptávky ze strany advokátních kanceláří jsme také vytvořili specifický advokátní slovník, který od roku 2010 také prodáváme.

Do jakých oblastí byste si představoval, že by měl v budoucnu NovaVoice dále pronikat nebo kde může být dále užitečný?

Každá oblast nasazení diktovacího softwaru má svá specifika, kterým NovaVoice postupně přizpůsobujeme tak, aby co nejlépe vyhovoval praxi. Průběžně také doplňujeme do slovníku nová slova.



Vladimír Moravec

Jaké je v současnosti procento úspěšnosti při rozpoznávání řeči? Je program použitelný, například pokud má diktující vadu řeči?

Úspěšnost rozpoznání se u specializovaných slovníků pohybuje v rozsahu 96 až 98 procent. Úspěšnost je závislá na schopnosti uživatele diktovat, osvojení si diktovacích pravidel pro formátování textu, číslovek, zkratek atd. S programem lze pracovat po instalaci a krátkém úvodním školení. Pokud diktující nemá vadu řeči, může začít ihned diktovat bez nutnosti adaptace na svůj hlas. Diktujícím s vadou řeči přizpůsobíme software tak, že nadiktuje připravený text, a my tím NovaVoice „naučíme“ jeho výslovnost.

Lze nějak vyjádřit časovou, eventuálně vycílit ekonomickou úsporu pro uživatele produktu?

Diktování je zpravidla tři až pětkrát rychlejší než zápis pomocí klávesnice. Výše úspory je dána složitostí stávajícího procesu přepisu. Jako příklad mohou uvést lékařské pracoviště, na kterém řešili nedostatek odborného administrativního personálu pro přepis historických nálezů. Nasazením NovaVoice nejenže nebylo nutné přijmout nové pracovníky, ale došlo i ke zkrácení doby vyhotovení a odeslání nálezů, protože lékař u mikroskopu nadiktovan nález přímo do informačního systému, jen ho zkontroloval a uložil. Podobný efekt časové úspory jsme zaznamenali u soudu. Soudce nadiktoval text, zkontroloval ho a asistentce předal už hotové usnesení k odeslání bez nutnosti další kontroly.

Je možné vyplňovat pomocí vašeho programu například předem připravené formuláře?

Náš program nahrazuje klávesnici a vepisuje text tam, kam umístíte kurzor, tedy i do formuláře.

Umožňuje program vyhledávat slova ve zvukových archívech?

To umíme. Pokud jsou zvukové archívy v dobré kvalitě, provedeme jejich indexaci, a pak je v nich možné vyhledávat podle klíčových slov podobně jako v textových dokumentech, přičemž rychlost vyhledávání je srovnatelná.

Umí program i opačně převést psané slovo do zvukové podoby?

Ano. Takovému převodu říkáme syntéza. Kvalita syntetického hlasu dosahuje úrovně lidského hlasu. Umíme simulovat třeba i nádech řečníka, a tím dosáhnout skutečně věrného přednesu.

Jaký další technologický vývoj svého produktu očekáváte – tedy, kam se bude dále posouvat, co očekáváte, že by v nějakém časovém horizontu mohl nového umět?

Produkt se neustále vyvíjí, zvyšujeme přesnost rozpoznání při zachování stejných nároků na výpočetní výkon. Do budoucna chceme uspokojit i běžného uživatele, který chce nadiktovat třeba e-mail nebo diplomovou práci. Pracujeme na slovníku s univerzální slovní zásobou, který umožní přepis mluveného slova se stejnou přesností, jako vykazují specializované slovníky, a přitom bude fungovat na běžném počítači.

Počítače se brzy naučí rozumět hovorové řeči

říká ředitel firmy, jež stojí za vývojem systému pro rozpoznání řeči
NEWTON Dictate

David Povolný

Společnost Newton Technologies nabízí ve spolupráci s Technickou univerzitou v Liberci systém automatického přepisu mluvené řeči do počítače – program Newton Dictate. Software je již delší dobu na trhu a podle ředitele společnosti Petra Heriana je v něm zapracována tříletá zpětná vazba velké řady uživatelů. Vývojový tým společnosti má podle něj bohaté zkušenosti se zaškolením soudců a státních zástupců v celém Česku.

Jaká je historie vašeho produktu?

Kdybychom měli zabrousit do samotných počátků, museli bychom se vrátit do roku 1993. Tehdy se totiž vědci z Ústavu informačních technologií a elektroniky Technické univerzity v Liberci pod vedením profesora Jana Nouzy začali soustavně věnovat výzkumu hlasových technologií. Výsledek jejich úsilí dostal časem podobu několika velmi úspěšných programů, které umožňují ovládat počítač, popřípadě pořizovat přepis mluveného slova jen s použitím lidského hlasu.

Jaké to byly programy?

Začalo to programem MyVoice, u kterého lidský hlas zcela nahradí klávesnici a myš, následoval MyDictate umožňující diktovat text po jednotlivých slovech a pak už konečně kontextuální diktovací program NEWTON Dictate, o kterém je řeč. Na jeho vývoji, jak už jméno napovídá, se podílela společnost NEWTON Media a od roku 2008 nově založená společnost NEWTON Technologies. Už v listopadu téhož roku jsme získali svého prvního klienta.

V jakých oblastech nachází váš software nejčastěji uplatnění?

Obecně je NEWTON Dictate velmi vhodný pro všechny profese, kde člověk potřebuje bezprostředně pořídit písemný záznam svého sdělení. Z logiky věci se přesto jako první nabízejí soudci, ale také státní zástupci či advokáti. V oblasti justice se NEWTON Dictate již dnes postupně stává důležitým a nepostradatelným instrumentem při psaní různých dokumentů. Strmě narůstá počet lékařů – chirurgů, radiologů, patologů a jiných –, kteří si pochvalují, kolik času jim náš přepisovací program ušetří. NEWTON Dictate je bezesporu jednou z cest, jak ulehčit práci také státní správě.

Do jakých oblastí byste si představoval, že by měl v budoucnu NEWTON Dictate dále pronikat nebo kde může být dále užitečný?

V příštích letech se o tom, jak může náš program ulehčit práci, budou moci přesvědčit inspektoři z České inspekce životního prostředí. V současnosti tyto inspektoráty vybavujeme notebooky s naším přepisovacím programem. Přířímým diktováním služebních zápisů jim neušetří jen čas, ale leckdy i nervy, když musí luštit své v rychlosti porízené poznámky. Svě uplatnění má NEWTON Dictate i v businessu a je velmi užitečným pomocníkem rovněž pro osoby s těžkým postižením horních končetin.

Jaké je v současnosti procento úspěšnosti při rozpoznávání řeči? Je program použitelný, například pokud má diktující vadu řeči?

Náš program umožňuje přepsat plynulě mluvené slovo s úspěšností až 98 procent. Přesnost přepisu mohou samozřejmě negativně ovlivnit faktory, jako je například špatná artikulace mluvěcího, ale i té se dá hlasová technologie uzpůsobit.



Petr Herian

Lze nějak vyjádřit časovou, eventuálně vyčíslit ekonomickou úsporu pro uživatele produktu?

Při monitoringu rozhlasových a televizních pořadů jsme si spolehlivě ověřili, že jejich klasický přepis trvá v průměru pětinašobek času v porovnání se skutečnou délkou nahrávky. Zato při úspěšnosti hlasové technologie vyšší než 90 procent pořídíme přepis hodinového zvukového záznamu za pouhé dvě a půl hodiny – v čistopisu, po korekturách.

Je možné vyplňovat pomocí vašeho programu například předem připravené formuláře?

Ano. Naše aplikace NEWTON Dictate to lehce zvládne. Lze ji používat jako hlasovou klávesnici, která píše na pozici kurzoru do libovolného programu.

Umožňuje program vyhledávat slova ve zvukových archivech?

S použitím naší technologie pro rozpoznávání hlasu můžete vyhledávat klíčová slova v audio i videoarchivech bez nutnosti pořizování textového záznamu. Poradí si přitom i s významně nižší úspěšností rozpoznávání, jelikož nejčastěji vyhledávaná klíčová slova jsou zpravidla snadno rozpoznatelná.

Umí program i opačně převést psané slovo do zvukové podoby?

Dnes jsou k dispozici různé programy pro převod psaného slova do zvukové podoby, některé najdete i jako freeware software na internetu. Hlasová syntéza je jednodušší úlohou, které se věnujeme jen okrajově. V její integraci do programu NEWTON Dictate nevidíme praktický přínos.

Jaký další technologický vývoj svého produktu očekáváte – tedy, kam se bude dále posouvat, co očekáváte, že by v nějakém časovém horizontu mohl nového umět?

Hlasová komunikace je přirozenou formou dorozumívání. Jsem přesvědčen,

že jednou se veškerá počítačová komunikace bude odehrávat jen na bázi hlasu, bez nutnosti pořizovat textové dokumenty. V dohlednější době se počítače nepochybně naučí rozumět hovorové řeči a zpracovávat dialogy či konverzace více mluvčích. V krátkodobé prognóze pro příští rok pak očekávám, že většina čtenářů vašeho časopisu si už svou práci bez diktování nebude umět představit (*směje se*).