

# NEWTON Dictate

## Úvodní školení testovací skupiny

**Oblastní inspektorát Brno**  
**23. listopadu 2010**



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST

[www.esfcr.cz](http://www.esfcr.cz)

## Obsah školení

### I. část (2 hodiny)

- představení NEWTON Technologies a školícího týmu,
- nasazení hlasových technologií na ČIŽP, co je cílem testovací skupiny
- školení – program NEWTON Dictate,
- školení „jak správně diktovat“

### *Přestávka*

### II. část (2 hodiny)

- uživatelská podpora – telefonní a e-mailový helpdesk pro uživatele ČIŽP
- otázky a diskuse
- individuální školení – diktování vzorových textů, problematická místa při diktátech
- dohodnutí termínu individuálních konzultací po 14 dnech, sběr dat od testovací skupiny

# I. část

## Představení NEWTON Technologies a školícího týmu



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)

## **NEWTON Technologies, a.s.**

- **2003 – 10/2008** projekt rozpoznávání hlasu v rámci NEWTON Media, a.s., implementace systému rozpoznávání plynulé řeči do procesu doslovného přepisu televizních a rozhlasových pořadů
- **27. 10. 2008** – založení společnosti NEWTON Technologies, a.s.
- **01 / 2009** – vítěz II. etapy pilotního projektu Ministerstva spravedlnosti (pronájem technologie do doby vypsání výběrového řízení na nákup diktovacích systémů)
- **11 / 2009** – vstup na slovenský trh s rozpoznáváním řeči pro lékařský obor radiologie
- **2010** – intenzivní sběr dat a rozvoj odborných slovníků
- **06 / 2010** – vítěz výběrového řízení pro Ministerstvo vnitra

## Sesterská společnost Newton Media, a.s.

- 14 let na trhu, 150 zaměstnanců, 3 země (Česká republika, Slovensko, Polsko), v ČR 1000 velkých klientů (Úřad vlády, Parlamenti, ministerstva, velké firmy, banky, nadace,...)
- **Největší komerční archiv plných textů v ČR** (větší než Seznam)
- Unikátní patentovaný systém zpracování dat, (světová konference FIBEP 2008 v Praze)
- Multimediální archiv pro vydavatele a vysílatele (Mafra, Ringier, TV Nova, ČRo, PSP,...)

## Spolupráce s Technickou univerzitou v Liberci (od roku 2005):

- Propojení vývojových týmů, sdílení výsledků a zkušeností, zpětná vazba
- Spolufinancování vývoje technologie ATT
- Samostatné nadstavbové moduly, nové aplikace, servis, komunikace s klienty

## NEWTON Technologies – tým pro uživatele ČIŽP

Ing. Petr Herian – koordinátor projektu, technik, lektor

Ing. Magdalena Horánska – odborný garant, technik, lektor

Ing. Stanislav Čihák – technik, lektor (tvorba slovníků)

Pavel Barták – technik, metodik, lektor

Jan Bechyňák – technik, metodik, lektor

Mgr. Kristýna Srbová – metodik, evaluátor, lektor

Karolína Perlová – metodik, evaluátor, lektor

Magdaléna Hladká – metodik, evaluátor, lektor

Petr Pazour – technik, lektor, integrace hlasových technologií do systémů ČIŽP

Mgr. Jan Šváb – technik, lektor, integrace hlasových technologií do systémů ČIŽP

= CELKEM 7 techniků a 10 lektorů

**Ministerstvo spravedlnosti** - projekt pro všechny soudy a státní zastupitelství v ČR (přes 1.600 licencí)

**Ministerstvo vnitra** – vybavení odboru Ministerstva diktovacím systémem

**Parlament ČR - Poslanecká sněmovna** - implementace systému přepisovacího pracoviště

**Česká národní banka** - implementace systému přepisovacího pracoviště

**NEWTON Media, a.s.** - projekt přepisu rozhlasového a televizního vysílání, titulování TV

Česká zemědělská univerzita

1. a 3. Lékařská fakulta Univerzita Karlova

Všeobecná fakultní nemocnice v Praze

IKEM - Institut Klinické a Experimentální Medicíny

Masarykova nemocnice v Ústí nad Labem

Nemocnice Prachatice, a.s

Mammacentrum, Praha

ACHK - Ambulantní centrum pro choroby hlavy a krku, s.r.o.

Gymnázium, ZŠ a MŠ pro sluchově postižené, Praha

Metropolitní univerzita Praha

Masarykova univerzita v Brně

Oblastní nemocnice Mladá Boleslav, a.s.

Poliklinika Budějovická

Poliklinika Prosek

Revmatologický ústav

MEDICON, a.s.

Manlomka, s.r.o., chráněná dílna



# I. část

**Nasazení hlasových technologií na ČIŽP,  
co je cílem testovací skupiny**



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)



## Projekt nasazení technologie pro rozpoznávání hlasu pro zaměstnance ČIŽP financovaný z prostředků EU

- realizována dodávka 140 licencí programu NEWTON Dictate vč. notebooků a headsetů
- I. fáze – testovací skupina, která má 30 uživatelů (Olomouc, Brno, Praha)
- 3 měsíce budou probíhat vždy po 14 dnech individuální konzultace a intenzivní sběr dat (nadiktované texty, nová slova, připomínky k fungování programu)
- na základě podnětů a připomínek připravíme program NEWTON Dictate přímo na míru pro širší využití v ČIŽP (II. fáze)

**= ČÍM VÍCE PODNĚTU OD VÁS ZÍSKÁME,  
TÍM VÍCE BUDE NEWTON DICTATE V ČIŽP POMÁHAT**

# I. část

## Školení – program NEWTON Dictate



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)

# Program NEWTON Dictate

- Významně zefektivní práci při zapisování delších textů.
- Šetří čas i náklady.
- Přepisuje mluvené slovo v reálném čase s přesností přes 95 % (u zkušeného uživatele)
- Funguje pro jakýkoliv (zdravý) hlas, adaptace na uživatele je možná, ale není nutná.
- Do slovníku lze jednoduše vkládat vlastní slova, sousloví i šablony.
- Umožňuje opravy a editaci textu.
- Funguje i jako elektronický diktafon.
- Automaticky formátuje čísla, symboly i interpunkci.
- Je k dispozici v několika verzích: NEWTON Dictate (samostatná diktovací aplikace), NEWTON Dictate Bar (verze určená pro zapisování na pozici kurzoru s možností fixace), NEWTON Dictate Offline (serverové řešení)

# Školení SW NEWTON Dictate

- hlasové profily
- ukázka diktování
- přehrávání zvukové stopy
- práce s textem a zvukovou nahrávkou (dodiktování, korektury, ukládání, exporty, formátování textu)
- jak přidat slovo do slovníku, práce se šablonami, práce s uživatelským slovníkem
- jak přizpůsobit program na vlastní hlas
- možnosti ovládání hlasem
- rozpoznání nahrávek pořízených digitálním diktafonem
- vyplňování formulářů



esf evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)

## ŘEŠENÍ NA KLÍČ

- Možnost diktovat kdykoliv a kdekoliv
- Snadná a příjemná obsluha
- Integrace mikrofonů, diktafonů, nožních pedálů
- Možnost editace záznamu
- Propojení do PC nebo do sítě



# I. část

## Školení – „Jak správně diktovat“



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)



## Přepis hlasu - principy

### Složitost češtiny

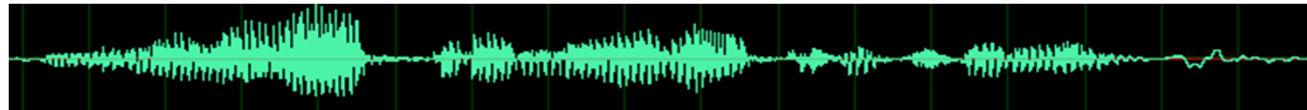
- **Ohebný jazyk** (angličtina ~ 50 000 slov, čeština > 1 000 000 slov a tvarů)
- **Volná struktura** věty, mnoho stejně znějících slov, 250 000 příjmení
- **Nelze převzít řešení** z jiných jazyků - Microsoft, IBM, Nuance, Philips

### System pro přímé diktování do PC

- Slovník – **320 000 nejfrekventovanějších slov a spojení** (možnost nadiktovat prakticky libovolný text).
- Zahrnutí slov a frází **typických pro oblast životního prostředí.**
- Automatické **formátování čísel a symbolů, interpunkce.**
- Rozpoznávání řeči funguje **pro jakýkoliv (zdravý) hlas**
- Text se na obrazovce objevuje prakticky **okamžitě**, takže jej lze očima sledovat a průběžně pokračovat v diktování.



Ukázka záznamu řeči:



***Co bylo řečeno a v jakých časových okamžicích?***

-	zavolá	-	premiéra	předseda	-	
-	za	vola	premiéra	předseda	-	
-	zvolá	prima	éra	před	sebou	-
-	zavolej	premiéra	předsedo	-		
-	zavolá	-	premiér	a	předseda	-

**Při rozpoznávání spojitě řeči není apriori známo:**

- Kolik slov bylo řečeno?
- Jaká sekvence slov byla řečena?
- Byla všechna vyřčená slova ze slovníku?
- Byla to skutečně jenom řeč nebo i další zvuky a hluky?
- V jakých časových okamžicích začínala jednotlivá slova?

Komunikace řečí - pro člověka přirozená věc,  
pro počítač velmi komplikovaná úloha

# System pro rozpoznávání hlasu se stále učí

1. HLAS - Variabilita hlasového projevu, každý člověk mluví jinak
2. SLOVNÍK - Rozsáhlost slovní zásoby
3. KONTEXT - Volba nejpravděpodobnější sekvence po sobě jdoucích slov (odlišná témata)
4. PROSTŘEDÍ - Vliv prostředí a přenosové cesty - degradace signálu



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)

## Uživatelské desatero

1. Diktují se čárky, tečky a další **interpunkční znaménka**.
2. Pro diktování jednotlivých písmen použijte znakovou abecedu.
3. Snažte se **diktovat souvisle** v rámci věty.
4. Systém rozpoznává lépe **delší fráze**, používejte nezkrácené tvary slov.
5. Hlasovou adaptaci provádějte až poté, co si osvojíte pravidla diktování.
6. Mikrofon mějte umístěný **asi 5 cm před ústy**, spíše při kraji úst či pod ústy.
7. Pracujete-li s notebookem, mějte jej zapojený do elektrické sítě.
8. Do programu **můžete přidávat nová slova**, věty, odstavce i libovolně dlouhý text pomocí tzv. hlasových maker.
9. Editovat a opravovat text je možné i během poslechu diktátu.
10. Program má automatické ukládání. Nastavte vhodný interval ukládání nebo automatické ukládání podle potřeby vypněte.

## Jak diktovat...?

- zkratky (obecná pravidla, manuál – výslovnost zkratek ve státní správě)
- diktování speciálních symbolů (nový odstavec, uvozovky atp.)
- diktování čísel a odstavců
- diktování peněžních částek a časových údajů
- tituly, hodnosti, obchodní společnosti
- zákony, paragrafy, odstavce a písmena
- spisové značky a čísla jednací
- použití hláskovací abecedy

# Přestávka



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)

## II. část

### Uživatelská podpora



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)

# HELPDESK PRO UŽIVATELE ČIŽP

Byl zřízen e-mailový helpdesk

[cizp@newtontech.cz](mailto:cizp@newtontech.cz)

Telefonický helpdesk 24 hodin denně

225 540 225

725 540 225



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)



## II. část

**Vaše otázky, diskuse**



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)

## II. část

### Individuální školení

# Individuální nácvik diktování s lektory

- diktování vzorových textů
- nácvik diktování problematických výrazů
- nácvik přidávání slov a šablon do slovníku
- přehrávání a export

## II. část

**Dohodnutí termínu příští individuální konzultace**

# Individuální konzultace po 14 dnech

- doškolování
- sběr nadiktovaných textů ve formátu .tta (je potřeba ukládat)
- sběr přidaných hlasových maker
- připomínky a podněty

**Děkujeme za pozornost a  
těšíme se na naši spolupráci**

**Ing. Magdalena Horánska**

**Jan Bechyňák**

**Mgr. Kristýna Srbová**



evropský  
sociální  
fond v ČR



EVROPSKÁ UNIE



OPERAČNÍ PROGRAM  
LIDSKÉ ZDROJE  
A ZAMĚSTNANOST

PODPORUJEME  
VAŠI BUDOUCNOST  
[www.esfcr.cz](http://www.esfcr.cz)