

## ● Problém, ktorý firmy reálne majú

Firmy pracujú s veľkým objemom slovenského textu:

- dokumenty, e-maily, zmluvy, komunikácia
- fulltextové vyhľadávanie často zlyháva kvôli skloňovaniu a časovaniu
- rôzne tvary slov sa nezhodujú → nepresné výsledky
- text bez diakritiky (e-maily, chat) ešte zhoršuje situáciu

👉 Výsledok:

- horšie vyhľadávanie
- nižšia kvalita analýzy dát
- vyššie náklady na manuálnu prácu

## ✓ Naše riešenie

**LALS (Lexical Analysis Lemmatization Service)** je vysokovýkonná služba pre spracovanie slovenského textu. Rieši problém slovenskej morfológie, ktorý bežné vyhľadávanie nezvláda.

👉 Prevedie všetky tvary slov na základný tvar (lemma)

👉 Jednoduché REST API → okamžitá integrácia

## ⚡ Kľúčové benefity

- 🇸🇰 Navrhnuté špeciálne pre slovenský jazyk
- 🔍 Výrazne zlepšuje kvalitu vyhľadávania
- ⚡ Latencia v jednotkách milisekúnd (aj pri záťaži)
- 🚀 Viac ako **400 000 slov / sekundu**
- 🧠 Podpora všetkých slovných druhov, vrátane priezvisk, ktoré majú v slovenskom jazyku odlišné skloňovanie ako podstatné mená
- 📄 Funguje aj pre text bez diakritiky
- 🛠️ Jednoduchá integrácia (REST API)
- 🏠 On-premise riešenie (žiadne dáta mimo firmy)

## 📊 Ukážka

**Vstup:** „Výročná správa o činnosti Slovenskej televízie“

**Lematizácia:** výročný správa o činnosť slovenský televízia

👉 Vhodné ako vstup pre vyhľadávanie, NLP a analýzu dát

## 🧠 Prečo lematizácia (nie stemming)?

Slovenčina má bohatú morfológiu – jednoduché „orezanie slov“ (stemming) nestačí. Rôzne tvary slov po stemmingu vedú k tomu, že vyhľadávanie nenájde všetky tvary slova.

**Príklad:**

**Slová:** človek, ľudia

**Slová:** íst', ide, išiel

✗ **Stemming:**

→ človek → člov

→ ľudia → ľud (úplne odlišné)

✗ **Stemming:**

→ íst' → ís

→ ide → ide

→ išiel → iš (neprepojené tvary)

✓ **Lematizácia:**

→ človek → človek

→ ľudia → človek

✓ **Lematizácia:**

→ íst' → íst'

→ ide → íst'

→ išiel → íst'

## 📌 Príklady použitia

Typické situácie, kde LALS prináša hodnotu:

- 🔍 fulltextové vyhľadávanie (Solr, Elasticsearch)
- 📄 spracovanie dokumentov (DMS, archívy)
- 📧 analýza e-mailov a komunikácie
- ⚖️ právne a compliance systémy
- 🏛️ eGovernment systémy
- 🤖 NLP pipeline (predspracovanie textu pre AI / analýzu)

## 🧩 Rozšírenia (upsell potenciál)

Možnosti rozšírenia:

- intelligent query expansion (inteligentné rozšírenie dotazov o synonymá a významovo prepojené výrazy)

## 🤝 Príležitosť pre partnera

Hľadám partnerov (integrátori, IT firmy), ktorí chcú:

- integrovať LALS do svojich riešení
- zlepšiť kvalitu vyhľadávania u klientov
- predávať ako súčasť vlastného produktu
- bez potreby vlastného vývoja NLP riešenia

👉 Ideálne pre:

- DMS systémy
- enterprise search riešenia
- vlastné SaaS / on-premise produkty

## 💰 Obchodný model pre partnera (orientačný)

Položka	Cena	Partner
Licencia	5 000 – 15 000 €	50 %
Integrácia	3 000 – 10 000 €	100 %
Údržba	1 000 – 5 000 €/rok	50 %

👉 Partner zarába najmä na integrácii a službách

## 🚀 Demo

Ponúkam krátke demo (5–10 minút):

- reálne spracovanie slovenského textu
- porovnanie výsledkov vyhľadávania
- výkon v praxi (latencia, priepustnosť)
- jednoduchá integrácia

## ✉️ Kontakt

**[Róbert Baláž]**

[bohem303@gmail.com]

[0950 259 480]

[https://bob303.duckdns.org]