

# Systém na automatizované doporučování obsahu

Český rozhlas hodlá v rámci projektu mujRozhlas, audioportálu s archivem vysílání ČRo, živým vysíláním a podcasty, nabízet uživatelům audioobsahy na základě jejich zájmů a preferencí. Monitoring, analýzu chování a nabízení personalizované nabídky obsahu by měl zprostředkovávat doporučovací systém (recommendation/recommender system) a to ve všech předpokládaných kanálech jakými jsou aplikace pro mobil a tablet, webová aplikace, HbbTV či systémy pro automobily.

## Použité zkratky

ČRo - Český rozhlas

mR - mujROZHLAS

RS - recommendation system / doporučovací systém

Admin - administrátor na straně ČRo

Reálný čas - tím je míněno období, ve kterém uživatel nevnímá poskytnuté informace jako opožděně doručené

BE - back-end

FE - front-end

## 1 Základní popis systému

1. RS bude na základě chování uživatele navrhnout v reálném čase k další konzumaci obsahy, o které se zajímá.
2. ČRo předpokládá, že dodávaný RS bude hybridním systémem běžných technik (kolaborativní filtrování, obsahová analýza / filtrování) a vlastního řešení dodavatele.
3. RS bude nabízet položky na základě chování uživatele a dle obsahu položek.
4. RS bude nabízet personalizovaný obsah / personalizovaný playlist, personalizované našeptávání a vyhledávání, a to na základě obsahu (obsahová analýza), chování (kolaborativní filtrování) a vytvářet tak (predikovat) atraktivní, doposud neposlouchaný mix obsahu ke konzumaci.
5. RS bude umožňovat uživateli základní customizace ve smyslu zakázání zobrazování určitých témat - ke zvážení.
6. Kódy RS jsou použitelné pro webové prostředí i pro mobilní aplikace či další aplikace (HbbTV, aplikace do automobilu...)
7. Admin bude mít možnost ručním zásahem vkládat do nabídky vlastní časově omezené položky / položku.
8. Admin bude mít možnost měnit váhy položek / témat, umožňuje-li to systém.
9. Admin bude mít přístup k souhrnným statistikám dle zadaného období.

## 2 Rozsah dodávky

1. Vlastní RS řešení pro BE provozované na HW dodavatele
2. Základní administrace přístupná uživatelům / administrátorům ČRo pod autentizací (např. pro nastavení, statistiky, váhy apod. - dle možností systému)
3. Dokumentace k implementaci a s tím související podpora.
  - a. Kódy pro FE (mobilní aplikace, web ... atd.), implementaci zařizuje ČRo a smluvní dodavatelé ve spolupráci s dodavatelem.

4. Dodavatel RS by měl podat vysvětlení, jakým způsobem jsou vytvářeny personalizované predikce, tedy další návrhy obsahů je konzumaci.
5. Dodavatel dodá API / rozhraní s dokumentací pro přístup k datům
6. Dodavatel dodá popis taktiky pro cold start.
7. Dodavatel popíše, jakým způsobem chce minimalizovat možný dopad uzavírání uživatele do uzavřené obsahové oblasti (bubliny).
8. Dodavatel dodá popis ochrany sbíraných dat o uživatelích.
9. Návrhy na další vylepšení systému pro mujRozhlas
10. Podpora vč. servisu po dobu trvání smlouvy.

### 3 Data - vstupní zdroje pro RS z ČRo

1. Metadata zhotovená na straně ČRo.
2. Ručně popsané audio (články z rozhlas.cz).
3. Automaticky popsané audio (AMS - strojová transkripce).
4. Automaticky vytvořené štítky (AMS + vyhledávání - strojové vytváření štítků).
5. Anonymizované chování a jednání uživatelů v aplikacích ČRo.

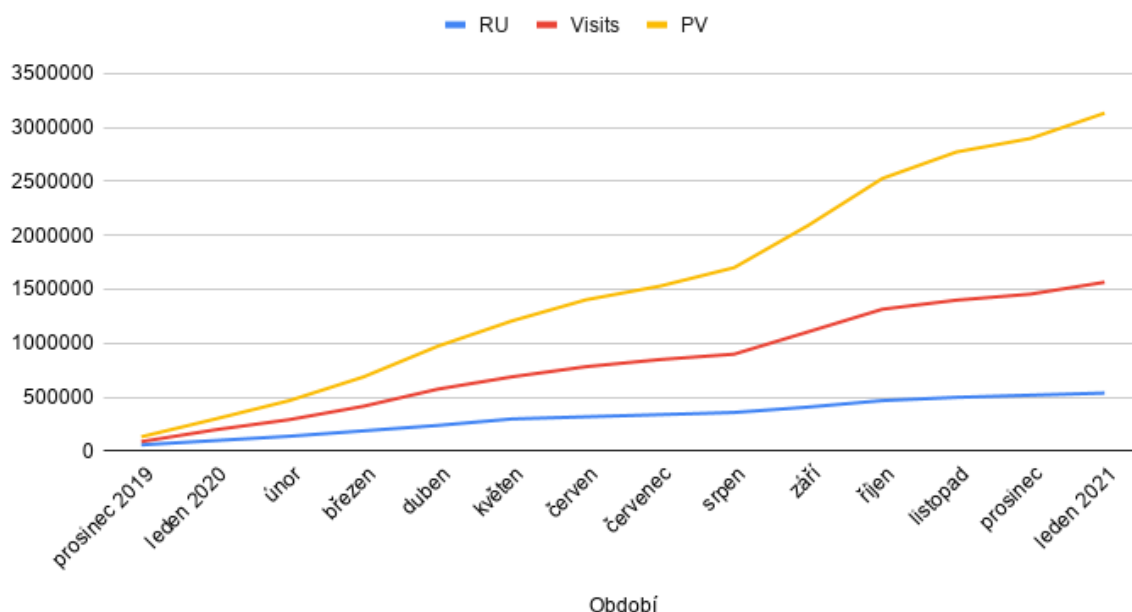
## 4 Podklady pro přípravu zakázky a výpočet ceny

### 4.1 Predikce vývoje návštěvnosti mR

Předpokládaná náročnost na RS dle vývoje počtu uživatelů

(<https://docs.google.com/spreadsheets/d/1tiJtL6ly0n-GUaSSB1rJUWK9r-CRBSZsfniile-Jw0M/edit?usp=sharing>)

Predikce vývoje návštěvnosti kanálů mR s doporučovacím systémem



## 4.2 Předpokládaný průběh zakázky po účinnosti smlouvy

### 4.2.1 Fáze

#### 1. fáze | Nastavení

Popis prací:

- Analytika stavu
- Konzultace
- Příprava a nastavení systému pro mR

Trvání: do 2 týdnů od účinnosti smlouvy

#### 2. fáze | Sběr dat pro testování služby a trénování algoritmů

Popis prací:

- Nasazení systému pro doporučování v reálném čase jen do (aplikace)
- Sběr dat
- Analytika interakcí

Spuštění: Následuje hned po ukončení fáze 1

Trvání: do 13 týdnů od spuštění fáze 2

#### 3. fáze | Příprava pro plné nasazení doporučování

Popis prací:

- Příprava doporučování obsahu v reálném čase na dalších kanálech projektu
- Příprava personalizovaných playlistů
- Příprava personalizovaného vyhledávání a našeptávání (ve spolupráci s našimi dodavateli pro vyhledávání a indexaci)
- Spuštění analýzy archivu ČRo a příprava AI modelů pro typy pořadů pro lepší vyhledávání a doporučování obsahu
- Příprava na cold start. Zvolení vhodné metody (nebo jejich kombinace) pro cold start a vhodného algoritmu (nebo jejich kombinací).

Spuštění: Nejpozději v polovině fáze 2

Trvání: nejpozději do 2 týdnů po skončení 2. fáze

#### 4. fáze | Nasazení základních funkcionalit

Popis prací:

- Nasazení personalizovaného doporučení v reálném čase
- Nasazení personalizovaných playlistů
- Nasazení aplikací pro cold start

Spuštění: Ihned po ukončení fáze 3

Trvání: Do ukončení smlouvy

#### 5. fáze | Inverzní doporučování

Popis prací:

- Nasazení inverzního doporučování - personalizované notifikace uživatelům

Spuštění: 4 týdny po ukončení fáze 3

Trvání: Do ukončení smlouvy

## 6. fáze | Vylepšování systému a případné návrhy na další vývoj

Popis prací:

- Správa systému
- Další ladění systému
- Další návrhy na vývoj a rozvoj

Spuštění: Ihned po spuštění fáze 5

Trvání: Do ukončení smlouvy

### 4.3 Předpokládaný harmonogram (critical path)

Harmonogram fází (prací) zavádění RS																																	
	Týden od podepsání smlouvy																																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	...	100	
Fáze																																	
1																																	
2																																	
3																																	
4																																	
5																																	
6																																	

## 5 Používání dat

- Dodavatel data o uživatelích a jejich interakcích nesmí předávat třetí straně bez souhlasu zadavatele.
- Po ukončení účinnosti smlouvy musí data uživatelů a jejich interakcí, personalizovaných playlistů apod. do 90 dnů dodat objednateli a následně odstranit ze systému.
- Data musí dodavatel zpřístupnit v elektronické podobě pro jejich uchování na straně zadavatele.
- Data musí být náležitě chráněna před zneužitím (dodavatel dodá hrubý popis zajištění).

## 6 HW

- HW pro RS je na strojích dodavatele a HW není instalován v ČRo.
- ČRo případně upraví dle možností požadovanou kapacitu linek mezi interními systémy.

## 7 Podpora

- Integrovaná podpora v rozsahu do 25 hodin měsíčně.
- Dostupnost v pracovních dnech 8-16.
- Konzultační služby v rozsahu 4 hodiny měsíčně.