

# Inteligentní systém pro analýzu a predikci veřejné dopravy

Kolektiv autorů řešitelů projektu



# Informace o projektu

## Základní identifikace projektu:

- Název: Inteligentní systém pro analýzu a predikci veřejné dopravy
- Ev. číslo: CK01000165
- Program: DOPRAVA 2020+
- Smluvní partner: Technologická agentura České republiky
- Garant: Ministerstvo dopravy České republiky, Odbor veřejné dopravy

## Řešitelé:

- **CEDA Maps, a.s.** - digitální mapování, online data a služby spojené s mobilitou na pozemních komunikacích, mobilitou jako službou, infrastrukturou, elektromobilitou, parkováním, indoor navigací aj.
- **T-MAPY spol. s r.o.** – softwarové řešení pro dispečerské řízení veřejné dopravy, aplikace pro veřejnost, geoanalytické systémy
- **XT-Card a.s.** - využití bezkontaktních a multifunkčních čipových karet pro platební a dopravní aplikace
- **ČVUT Praha** - Fakulta dopravní
- **VŠE Praha** - Fakulta informatiky a statistiky

# Schválené výstupy projektu

1. Datové úložiště pro sběr dat o veřejné dopravě
2. Modul pro analýzu a střednědobou predikci pro potřeby organizátorů veřejné dopravy
3. Modul pro analýzu a krátkodobou predikci pro plánování tras
4. Modul pro optimalizaci personifikovaného osobního profilu multimodálního plánovače s podporou AI
5. SW nástroj pro podporu operativního a strategického řízení VD na základě aktuální situace a predikce poptávky
6. Poloprovoz Inteligentní systém pro analýzu a predikci veřejné dopravy

# Přípravná fáze prediktivních modelů

1. Identifikováno přes 33 hlavních predikovaných veličin v 6 hlavních skupinách
  - např.: Počet volných parkovacích míst (% obsazenosti); Počet volných dobíjecích bodů (% obsazenosti); Cestovní doba pro cestujícího; Predikce změny O/D matice v návaznosti na kulturní/sportovní/... akce, rozvoj průmyslových zón a další
2. Vybrané pilotní predikované veličiny
  - Cestovní časy na úsecích sítě pozemních komunikací
  - Počty cestujících vystupujících nebo nastupujících dle tarifní skupiny
  - Zpoždění na síti příměstské veřejné dopravy oproti jízdnímu řádu

Vysvětlující proměnné  
(atributy, vstupy,  
inputy...)

Vztah

Vysvětlovaná proměnná

Data

18	male	33.8		1	no	southeast
28	male	33.0		3	no	southeast
33	male	33.8		0	no	north
31	male	33.0		3	no	southeast
33	male		22.VII	0	no	northwest
32	male		28.IX	0	no	northwest
31	female		25.VII	0	no	southeast
46	female	33.4		1	no	southeast
37	female		27.VII	3	no	northwest
37	male		29.VIII	2	no	northwest
60	female		25.VIII	0	no	northwest
25	male		26.II	0	no	northwest
62	female		26.III	0	yes	southeast
23	male	34.4		0	no	southeast
56	female	39.8		0	no	southeast
27	male	42.1		0	yes	southeast
19	male		24.VI	1	no	southeast
52	female		30.VIII	1	no	northwest
23	male		23.VIII	0	no	northwest

Trénování

Model

4484.5  
725.55  
449.46  
1984.47  
3866.86  
3756.62  
240.59  
281.51  
406.41  
8923.14  
721.32  
7808.73  
826.84  
1090.72  
9611.76  
837.24  
10797.34  
395.17

Dotaz

10	female		27.09.2020		0	yes	southeast
18	male		33,8		1	no	southeast
28	male		33		3	no	southeast
33	male		22,07,2020		0	no	northwest
32	male		28,09,2020		0	no	northwest
31	female		25,07,2020		0	no	southeast
46	female		33,4		1	no	southeast

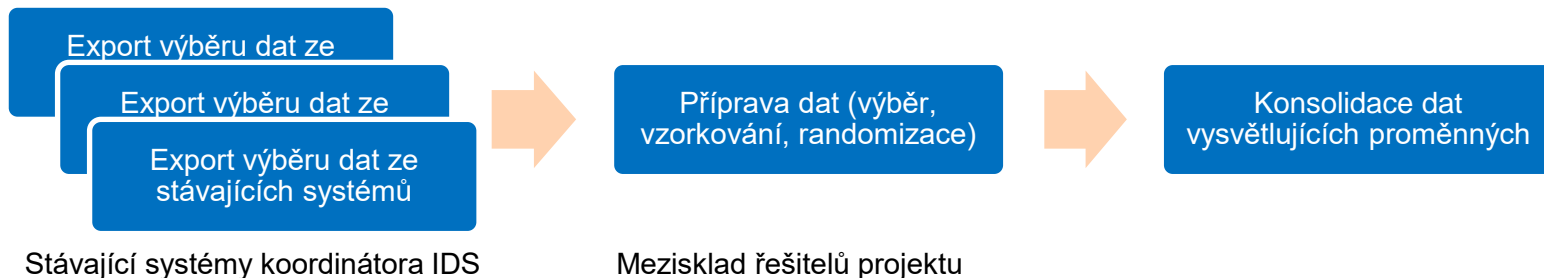
$Y \sim f(X_1, X_2, \dots)$

Predikce

4484.5  
725.55  
449.46  
1984.47  
3866.86  
3756.62  
240.59  
281.51

# Příprava dat pro predikční modely

- Napojení datových zdrojů (online) nebo získání offline
- Propojování dat
  - Lokalizace na síti komunikací
  - Sjednocení časových okamžiků pro měření
- Odstranění vadných dat
- Doplnění chybějících dat
- Harmonizace dat z různých zdrojů
- Export všech dat do 1 tabulky pro 1 predikční model
- Kódování proměnných pro použití v predikčních modelech



# Prediktivní model č.1

Vysvětlovaná proměnná  
Aktuální rychlost úseku  
komunikace

- Leader = CEDA
- Jednotka = Dopravně významné úseky sítě komunikací (36.850 TMC úseků)
- Krok = 1 hodina pro dlouhodobou predikci, 5-15 minut pro krátkodobou

## Vysvětlující proměnné

- Část dne, část týdne či roku, prázdniny (sezónní vlivy)
- Svátky, kulturní akce
- Počasí
- Charakteristiky segmentu úseku (např.: typ komunikace, třída komunikace, směr dopravního provozu, průměrná rychlost, průměrný sklon, nadmořská výška, rychlostní omezení, počet křižovatek a další)
- Mód dopravy
- Rychlost při plynulém provozu, data o zpoždění na síti komunikací z NDIC

# Prediktivní model č.2

## Vysvětlující proměnné

- Část dne
- Část roku, Prázdniny
- Svátky
- Kulturní akce
- Počasí
- Kapacita dopravního prostředku
- Začne se bez výluk, pak bude případně v krátkodobém modelu

Leader = XTC / T-Mapy

Jednotka = Zastávka& Spoj

Krok = den

## Vysvětlovaná proměnná

**Počet cestujících vystupujících nebo  
nastupujících (a jejich rozdělení  
podle tarifních skupin)**



# Prediktivní model č.3

- Predikce zpoždění spoje na mezi-zastávkovém úseku
- Snaha o predikci nezávislou na aktuálním jízdním řádu, pouze na charakteristikách provozu, typu vozidla, počasí, atd.
- Jednotka = mezi-zastávkový úsek
- Krok = hodina
- Využití při operativním řízení a optimalizaci příměstské dopravy

**Vysvětlovaná proměnná**  
**Odhadovaná doba projetí**  
**mezizastávkového úseku**

## **Vysvětlující proměnné**

- Vybrané vlastnosti z modelu č. 1 charakterizující aktuální propustnost, vztažené k mezi-zastávkovému úseku
- Základní charakteristiky mezi-zastávkového úseku
- Typologie zastávek
- Typologie vozidel
- Aktuální počasí

# Prediktivní model č.1- dlouhodobý a střednědobý

- Dlouhodobý a Střednědobí model
  - D-model vytváří dlouhodobé predikce (1-2 roky dopředu)
  - Část dne, část týdne či roku, prázdniny (sezónní vlivy, svátky)
  - S-model upravuje predikce z modelu Dlouhodobého o Počasí, kulturní akce

## DLOUHODOBÝ MODEL

Predikce Rychlost na úseku =  $f(\text{datumové proměnné, char. úseku}) + \text{náhodná chyba}$

## STŘEDNĚDOBÝ MODEL

Predikce náhodná chyba =  $f(\text{počasí, události})$

**Spojení predikcí: Predikce Rychlost na úseku + Predikce náhodná chyba**

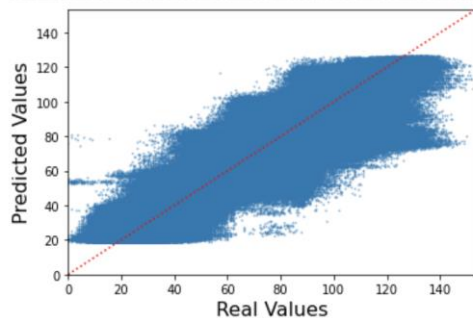
# Prediktivní model č.1- dlouhodobý a střednědobý

- Vyhodnocení modelu
  - Data jsou za období od **2020-03** do **2021-09**, na **1H granularitě**
  - Chronologicky posledních 20% pozorování je užito na testování modelu.

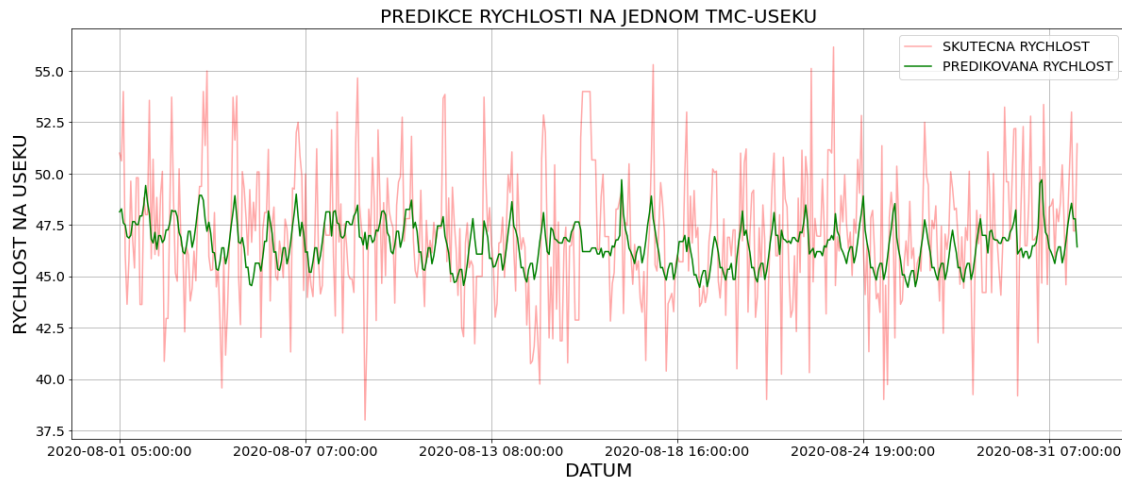
## Přes celý datový soubor

Metrics for regression:

```
=====
mean absolute error:      4.912
median absolute error:   3.480
mean squared error:      7.096
max error:               79.951
r2 score:                0.889
explained variance score: 0.890
Mean of True values:     54.705
STD of True values:      21.302
n:                       6256984.000
=====
```



## Přes jednotlivé úseky



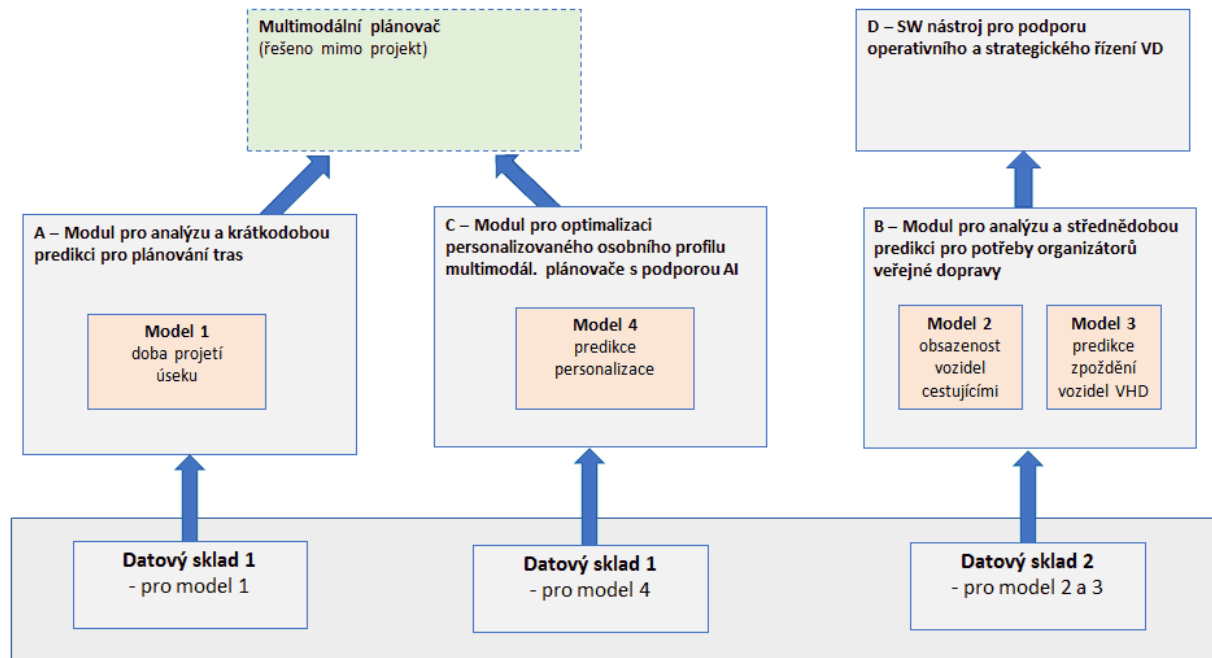
# Architektura řešení

1 datový sklad

4 predikční  
modely

3 SW moduly

1 SW aplikace  
(+ 1 mimo  
projekt pro  
demonstrace)



# SW nástroj pro podporu operativního a strategického řízení VD na základě aktuální situace a predikce

- Backend
  - Samostatný modul pro sběr, automatizované zpracování a ukládání dat z vozidel vč. historizace a možnosti zpřístupnění pro další využití
- Frontend
  - Rozhraní pro vizualizaci dat z ostatních SW modulů – webová (mapová) aplikace, reporting, geoanalytické výstupy
  - Možnosti přizpůsobení prostředí, připojování dalších datových zdrojů, individualizace obsahu výstupů
  - Možnost zakomponování výstupů do jiných aplikací

Aktualizace JŘ

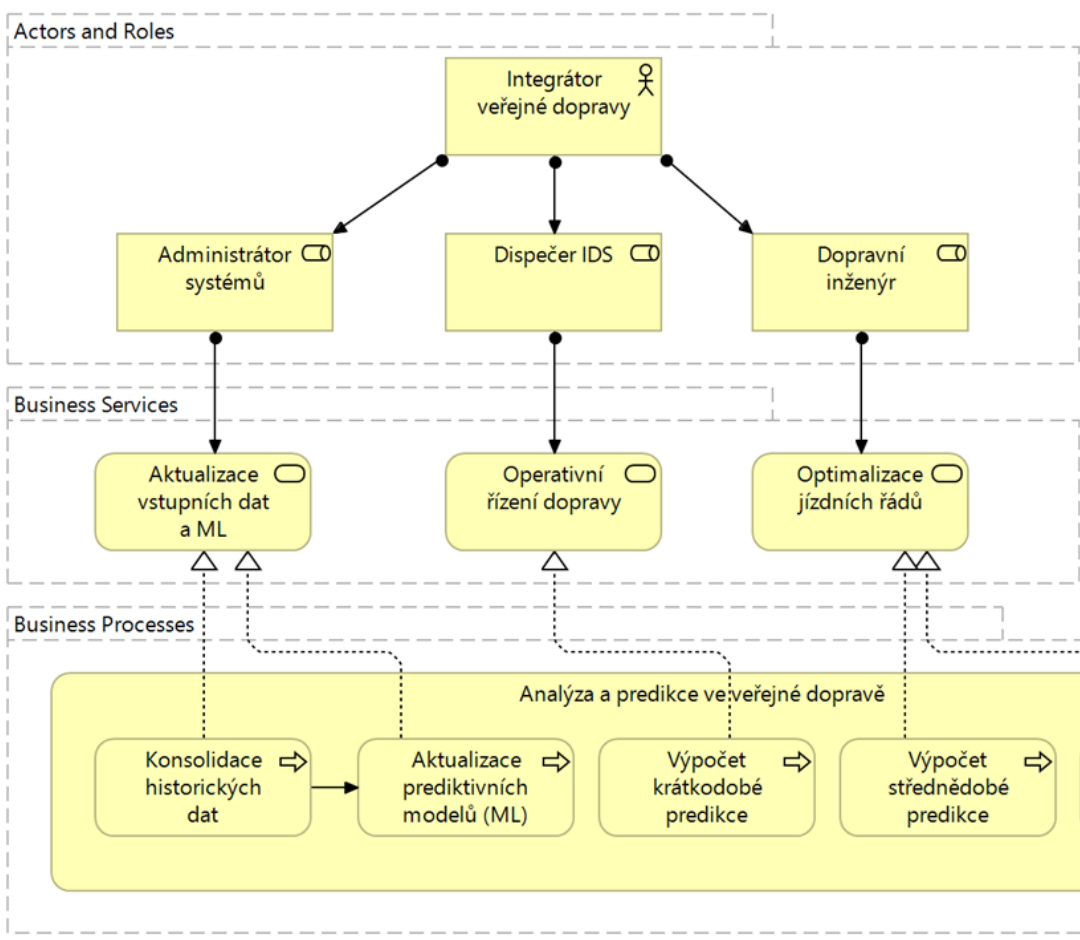
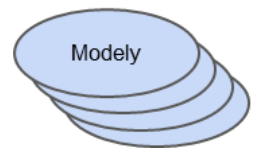
Příjem dat vozidel IDS

Zpracování historických dat

Převod TMC - MZÚ

Další backendové procesy

Služby poskytování výstupů modelů



Dispečink IDS

Samostatná mapová aplikace (Terinos)

Reporty

Obslužná aplikace

Mapová analytika

Volání modelů

Konsolidace dat, výpočty

# Integrace do systémů řízení VLD

- **Krátkodobé predikce**
  - Podpora práce dispečerů - notifikace ohrožených návaznosti spojů
  - Včasné informace pro cestující - zastávkové panely, aplikace pro veřejnost
- **Středně a dlouhodobé predikce**
  - Zajištění posilových spojů
  - Optimalizace jízdních řádů,
  - ...

The screenshot displays the IDS Dispatching - T-MAPY software interface. The top navigation bar includes 'Domů', 'Mapa', 'Kreslení', 'Odtazní', and 'Spoje'. The main window is divided into several sections:

- Detail:** Shows route information for '931 Kvasilny-Rychnov n.Kněž.-Vamberk-Ústí n.Orl.-Če. Výchukový' with contact details for AUDIS BUS s.r.o. and SPZ: 7H44726.
- STANICE:** A table listing stations with columns for 'Číslo', 'Km', 'Název', 'Plán', 'Skute', 'Zpo', 'Plán', 'Skute', 'Zpo', and 'Odh'. The row for 'Rychnov n.Kněž. - Pod Bu.' is highlighted in red.
- Map:** A map showing the route path through the region, with a scale of 1:5000.
- Návaznosti:** A table of connecting routes with columns for 'Spoj', 'Turnus', 'Číslo', 'Po', 'Název', 'Odjez', 'Zpožd', 'Čas', and 'Dopr'. It lists various bus lines and their destinations.
- Autobusové spoje (37):** A detailed schedule table with columns for 'Turn', 'Spoj', 'Z', 'Do', 'Odjezd', 'Příjezd', 'Zastávka', 'Poslední z', 'Přísti zpožd', 'Počet cest', 'Dopravec', 'SPZ', 'OBC', 'Výlukový', 'ID řidiče', 'Otrožená', 'Skupina', 'Příjezdy', 'MHD/VLD', 'Čas', and 'Aktu'. It provides a comprehensive overview of bus services.

# Děkujeme za pozornost