PRAHABOT 13

Cesta Prahy 13 k webovému AI asistentovi

Jakub Svoboda | Manažer kvality a kybernetické bezpečnosti, ÚMČ Praha 13



Fáze 1: Uzavřený pilot aneb digitální šanon

Zadání a prototyp (PoC):

Asistent pro sekci „Jak si zařídit“ (cca 80 článků napříč odbory).

Ruční sběr textů, čištění, vektorizace, prototyp ve Streamlit na subdoméně.

První bezpečnostní mantinely:

Záchyt vulgarit, zákaz dotazů na ohrožení života (přesměrování na tísňové linky IZS), GDPR.

Výsledek pilotního provozu:

Model ztrácel kontext a chyboval v češtině; malý vzorek neodhalil reálné dotazy.

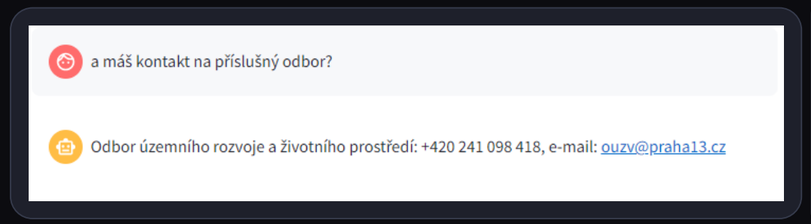
Rizika údržby: každá změna na úřadě = nutnost ručně přepisovat a nahrávat soubory.

Nečekaný bonus:

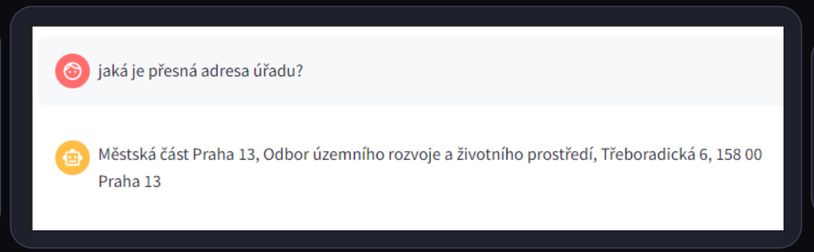
AI nastavila zrcadlo webu – odhalila zastaralé články a chybějící postupy.

Pilot: kontakt, adresa a ztracený kontext

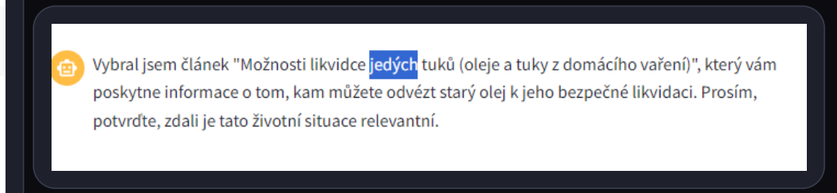
Kontakt na odbor



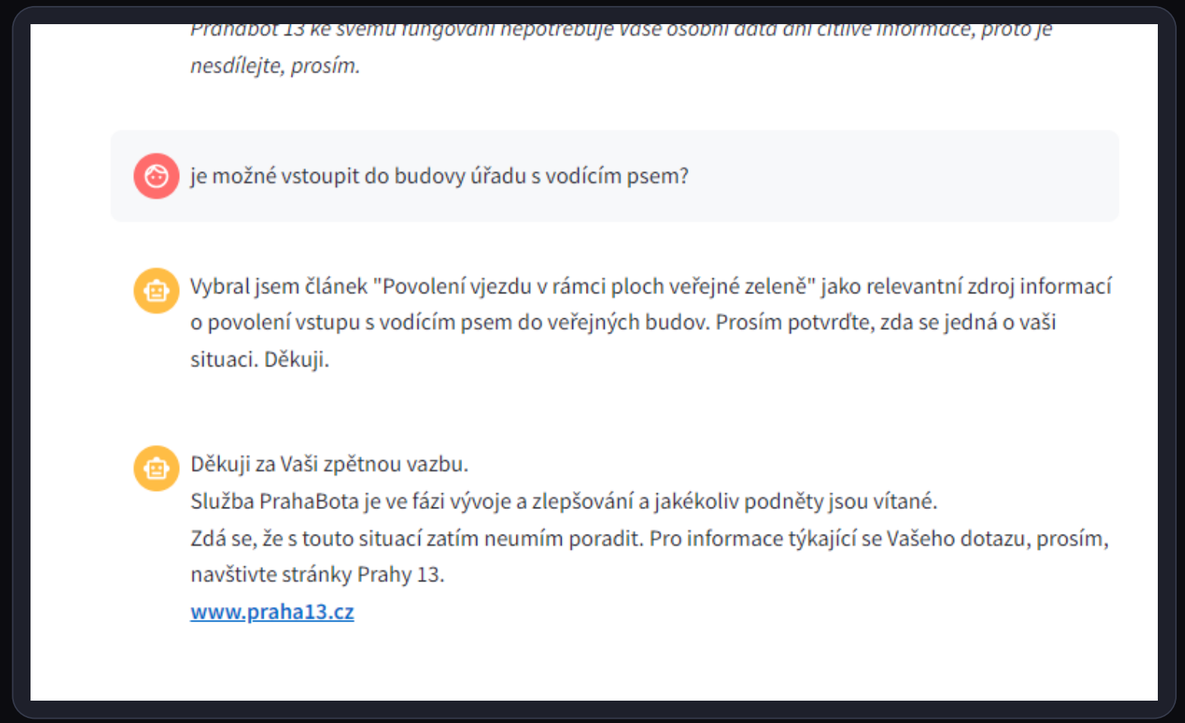
Adresa úřadu

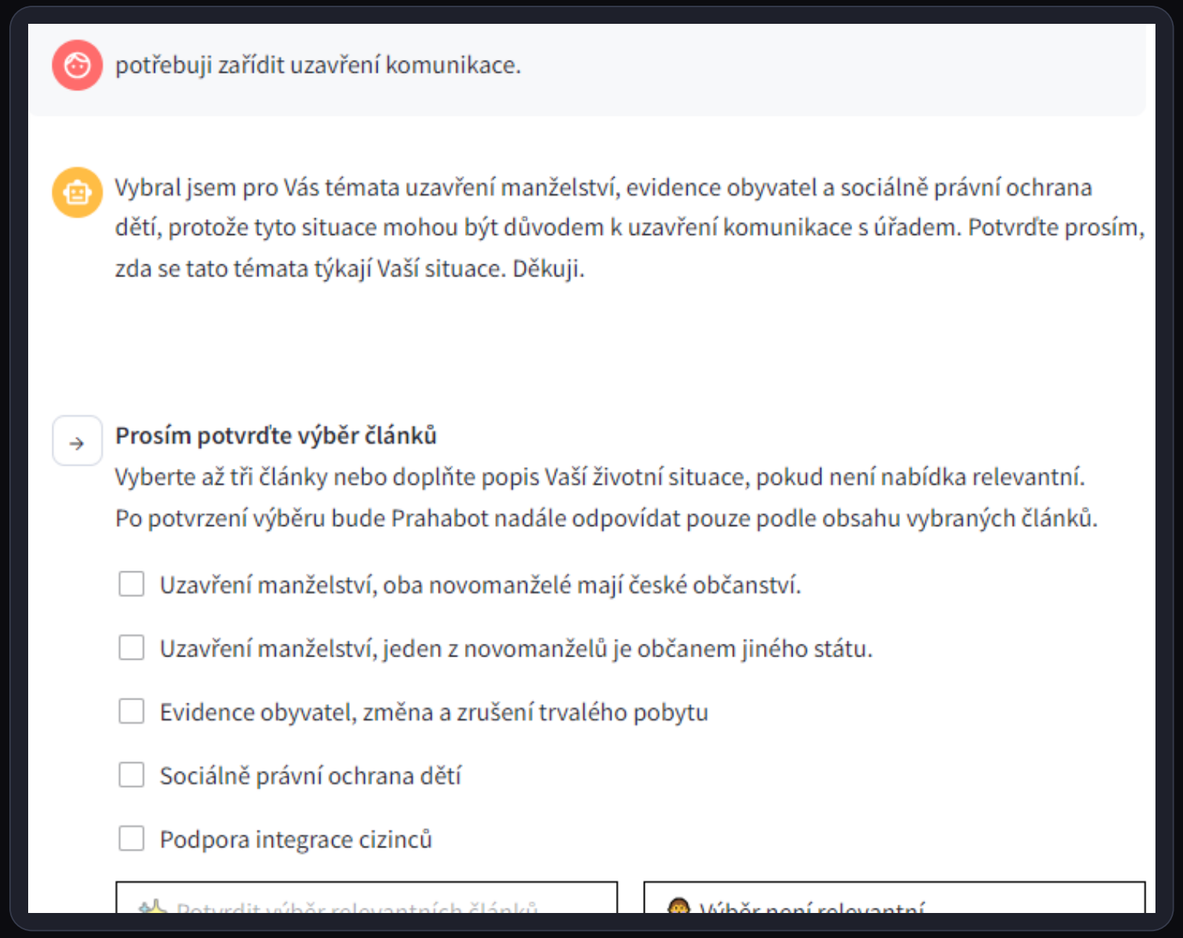


Ztracený kontext



Pilot: chyby při vyhledávání postupů





Fáze 2: Jak jsme (ne)napojili WordPress

Velká vize:

Centrální strukturovaná databáze znalostí celého úřadu.

Předpoklad doby: Pro budoucího telefonního voicebota na ústředně jsou nutná striktní tabulková data; scraping nestačí.

→

Technické řešení & střet s realitou:

Vývoj vlastních API endpointů napojených přímo do WordPressu.

Nulová technická dokumentace k roky bobtnajícímu webu.

Chaos v tagování a kategoriích → nutnost zdlouhavé ruční „digitální archeologie“.

→

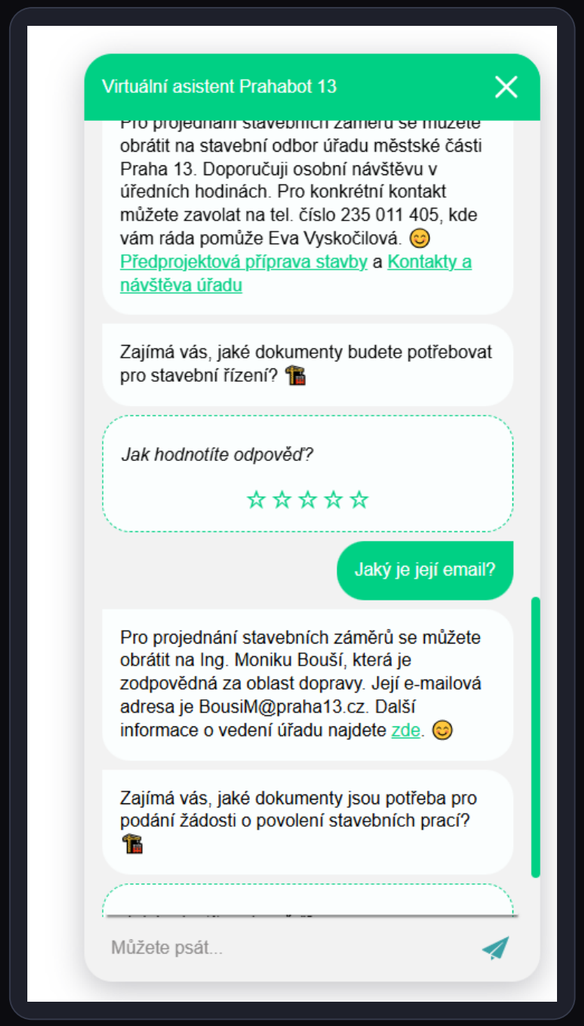
Výsledek a utopené náklady:

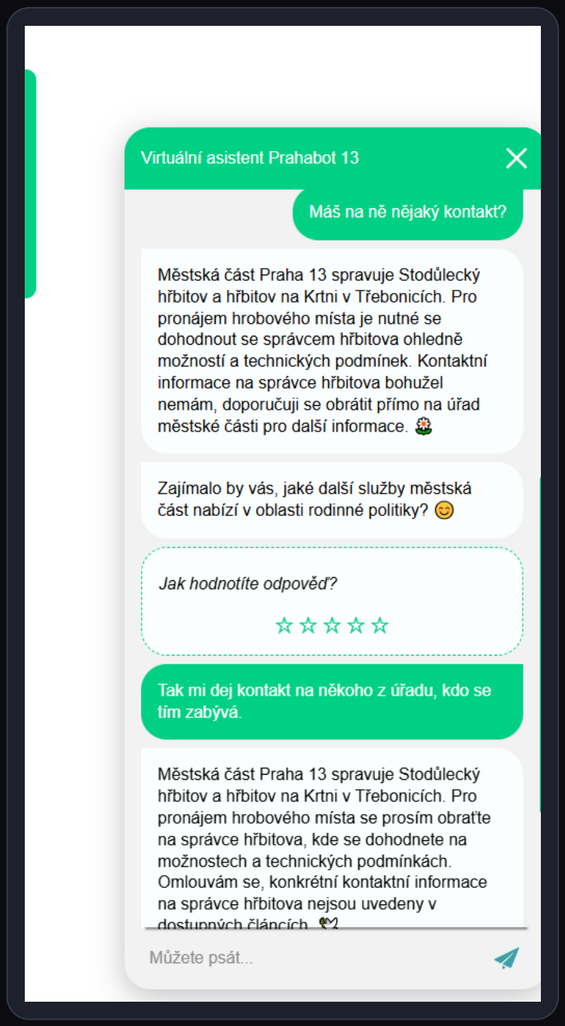
Vývojáři řešili výpadky API a třídění dat místo ladění bota.

Bot neznal základní úřední hodiny ani termíny voleb.

Slepá ulička: Měsíce testování bez výsledku a návrat k provizornímu lepení mezer přes PDF přílohy.

Chybějící kontakty a nesprávný kontext





Zlomový bod: Od slepé uličky k hotovému produktu

PŮVODNÍ CESTA

Vlastní vývoj (Build)

Krmení daty:

Bolestivé API na míru, chybějící dokumentace, věčné výpadky přenosu.

Práce s daty:

Požadavek na čisté tabulky a pevnou strukturu (slepá ulička pro web i voicebota).

Správa a lidské zdroje:

Tým úřadu v roli hasičů programátorských a databázových chyb.

SOUČASNÝ STAV

Hotový produkt (Buy)

Automatický crawling:

Robot čte web 1× denně (aktuality až 1×/h) – HTML, PDF, Word i úřední desku.

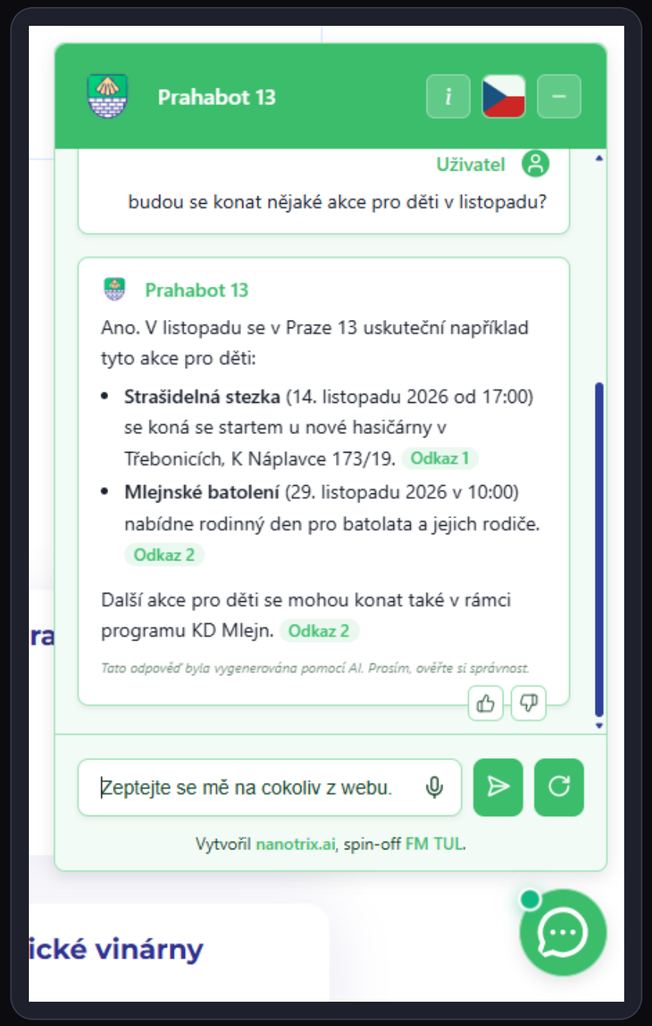
RAG architektura:

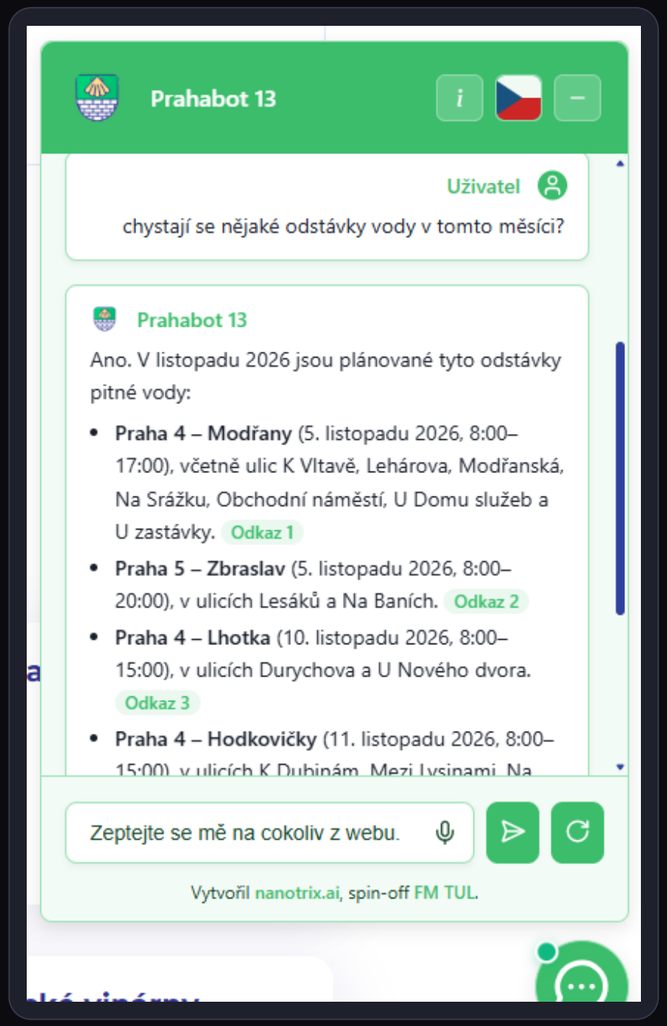
AI si sama najde odstavce v indexu a spolehlivě chápe běžný nestrukturovaný text.

Fokus na obsah a občana:

Infrastruktura běží v zabezpečeném klastru; úřad řeší kvalitu dat a práci s logy.

Hotový asistent: akce a odstávky vody





Realita provozu: Jak se lidé doopravdy ptají

Čísla z ostrého provozu: 1 694 dotazů za poslední 3 měsíce

88,5 %

věcných odpovědí

Úřední hodiny a kontakty (22 %), kultura a akce (12 %), pasy a OP (12 %), odpady (9 %), doprava (9 %).

49 %

dotazů mimo úřední dobu

28 % večer a v noci se špičkou ve 22:00, 21 % o víkendech. Služba 24/7 v praxi.

63 %

občanů přistupuje z mobilu

Mobilní úřad: přístup z mobilního telefonu.

Autentické situace z logů

„Kde si vyměním řidičák?“

Kompetenční paradox (radnice jako univerzální brána ke státu; bot musí odkazovat na Magistrát).

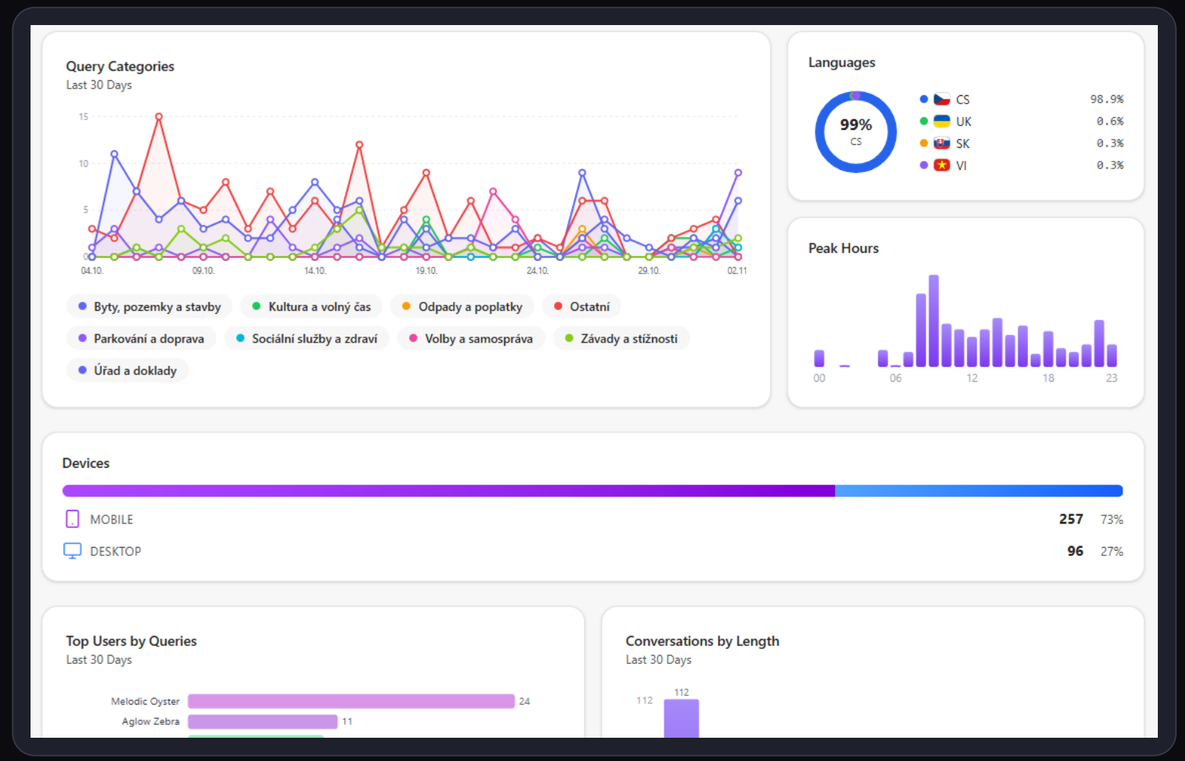
„Proč neteče teplá voda?“

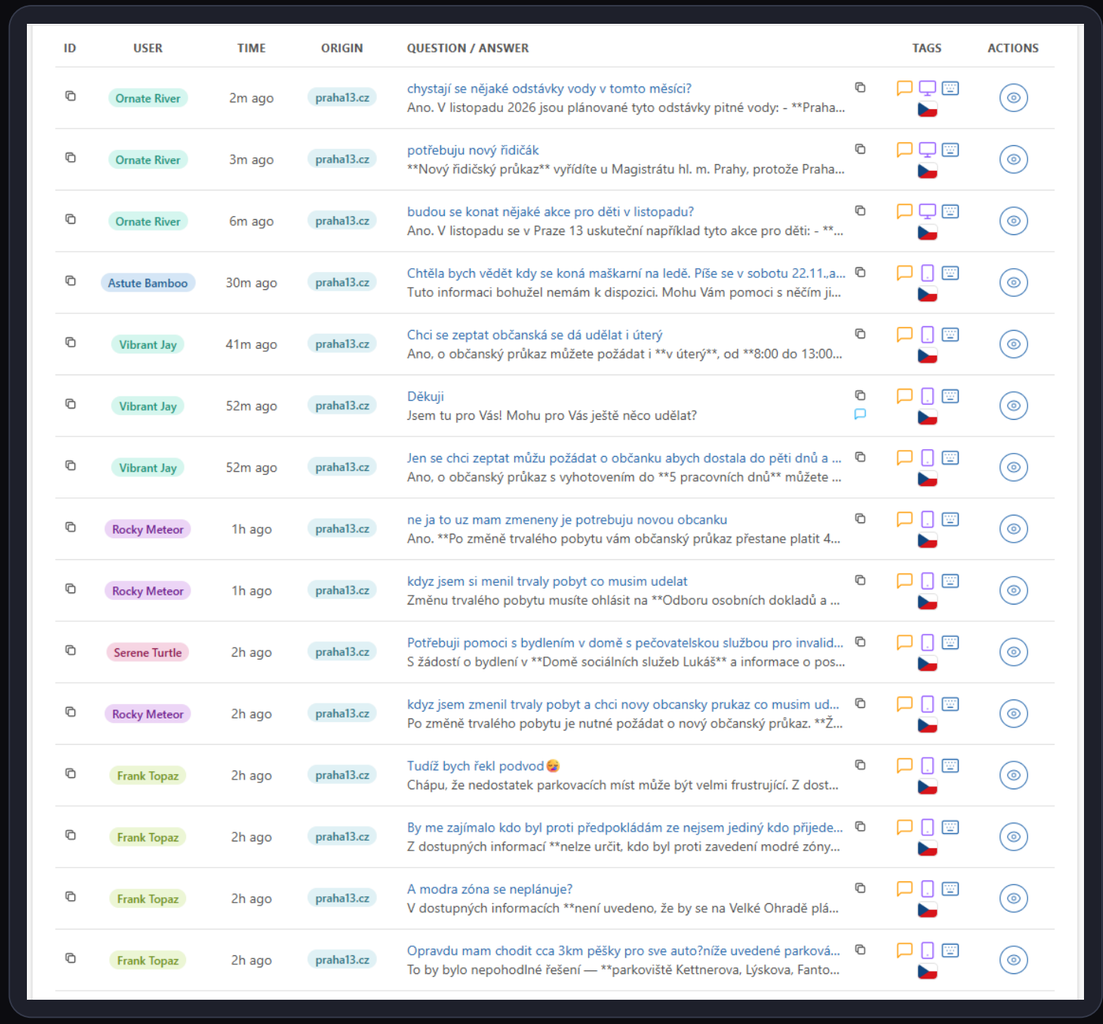
Radnice jako první instance (dotazy na havárie PVK, výpadky proudu a výkopy).

„Mám slevu na psa, když štěká jen v noci?“

Lidový folklór (pokusy usmlouvat poplatky či skládat básně o starostovi).

Dashboard a provozní přehled konverzací





Architektura v praxi: Jak asistent funguje pod kapotou

Krmení & indexace

Celý web MČ, ZŠ, KD Mlejn, měsíčník, usnesení, vyhlášky.

HTML, PDF i Office dokumenty.

Denní index (aktuality 1×/h).

Dotaz & hybridní hledání

Otázka občana (text / hlas).

Bezpečnostní filtr (útok/pozdrav).

Vektor dotazu + Full-text (nalezení přesných odstavců).

Generování odpovědi

LLM model (OpenAI / Google).

Odpovídá POUZE z nalezených dat.

Vždy uvádí ověřitelný odkaz.

Přirozeně v cizích jazycích.

Infrastruktura & soukromí

Provoz v zabezpečeném klastru (Kubernetes, Postgres, Python).
Nulový sběr osobních údajů: Asistent z principu nevyžaduje přihlašování, rodná čísla ani kontakty.

AI Governance

Řízení podle KPI (Dashboard)

Sledování trendů (zápisy do škol, volby, poplatky).

Fallback rate (12,3 %): Záchranná brzda = mapa bílých míst, co chybí na webu (např. chybějící euroklíče, storno rezervace).

Hodnocení palci využívá jen **1,4 % lidí** → klíčem k řízení je čtení textových logů.

Překlad lidštiny & trénování slovníku

Občan hledá „popelnice“, úřad píše „směsný komunální odpad“.

Správce propojuje hovorové dotazy s úřední terminologií.

Zkratky a úřední hantýrka: Občan nepíše celá slova. Logy ukázaly, že lidé masivně používají zkratky jako „OP“, „CP“ nebo „ZTP“. Bez ručního doplňování těchto zkratek do slovníku by AI u každého pátého dotazu tápala.

Bezpečnost, GDPR & AI Act

100 % odražených útoků (3,5 % provozu): Filtr zachytil všech 59 pokusů o prolomení. Zvládnutý koordinovaný nálet 1. 10. (42 útoků v base64, DAN, ADMIN OVERRIDE) i sociální inženýrství („záchranka pro babičku“).

GDPR paradox: Bot data nechce, ale lidé je sami píší → nutnost anonymizace logů a řízené retence.

Human-in-the-loop: AI potřebuje stálého lidského garanta.

5 tipů pro praxi, než spustíte vlastního bota

01 **Nestavte znovu kolo (Build vs. Buy):**
Vlastní vývoj jádra je dnes slepá ulička. Hotová SaaS služba s denním crawlingem ušetří roky trápení a statisíce z rozpočtu.

02 **Občan se ptá na život, ne na agendu:**
Nepočítejte s tím, že lidé znají kompetenční zákon. Naučte bota odkazovat i na magistrát, školy, dopravu a havárie.

03 **Pracujte s fallbacky:**
Kde bot neví, tam nemáte chybu v AI, ale díru v informacích na webu. Je to nejlevnější a nejpřesnější audit obsahu úřadu.

04 **Testujte mimo úřednickou bublinu:**
Zapojte seniory i veřejnost. Úředník trpí provozní slepotou (případně neochotou).

05 **Určete botovi lidského šéfa:**
Chatbot není bezúdržbový nástroj. Vyžaduje správce na čtení logů, bezpečnostní filtry, anonymizaci a soulad s legislativou.

Děkuji za pozornost

Úřad městské části Praha 13

Sluneční náměstí 2580/13
158 00 Praha 5 - Stodůlky

PRAHA
13

Jakub Svoboda

manažer kvality

235 011 650

svobodaj@praha13.cz