

VYSVĚTLENÍ ZADÁVACÍ DOKUMENTACE - SOUBOR DODATEČNÝCH INFORMACÍ ZADAVATELE Č. 2

Společnost: **Gitano cosmetics a.s.**
Se sídlem: Ruská 117/17, Pozorka, 417 03 Dubí
IČO: 06712363
Zastoupen: Vladimír Michalčík, člen správní rady

Název zakázky:

„Zvýšení digitální úrovně ve společnosti Gitano cosmetics a.s. – část 1 – AI voicebot“

Zadavatel poskytuje toto následující vysvětlení zadávací dokumentace vztahujícím se k výše uvedené zakázce.

Znění žádosti vysvětlení zadávací dokumentace č.1:

„1. Emoční a kontextová detekce v reálném čase

Požadavek na schopnost analyzovat minimálně tyto emoce: pozitivní, neutrální, negativní, hněv, radost, frustrace -- (Příl. č. 4a, bod 2 Funkční požadavky). Prosíme o upřesnění:

- o Jaký je zamýšlený praktický účel využití této funkce v obchodním procesu?*
- o Jak bude přesnost/míra správnosti vyhodnocení emocí měřena při akceptaci a s jakým minimálním prahem? o Jaké datové vzorky budou použity pro otestování - reálné hovory zadavatele vs. syntetické scénáře)?“*

Znění vysvětlení zadávací dokumentace č.1:

Cílem detekce emocí v reálném čase je zajistit, aby VoiceBot dokázal adekvátně reagovat na náladu volajícího a přizpůsobit tomu tón, styl i strukturu odpovědi. Prakticky to znamená např. zklidnění komunikace při detekci frustrace nebo přepojení na živého operátora při projevech hněvu.

Přesnost rozpoznání emocí bude vyhodnocována na základě testovacích scénářů simulujících reálné situace a interních nahrávek hovoru. Cílem je funkční spolehlivost a konzistentní chování VoiceBota, nikoli akademicky přesné skóre.

K ověření budou využity jak syntetické scénáře, tak anonymizované reálné hovory, které reflektují běžné provozní podmínky.

Znění žádosti vysvětlení zadávací dokumentace č.2:

„2. ISO/IEC 30122 – požadavek na soulad

Požadavek na soulad se standardem ISO/IEC 30122 – Voice command systems -- (Příl. č. 4a, bod 3 Technické požadavky). Prosíme o upřesnění:

- o Jaký je účel požadavku na soulad s tímto standardem – plánuje zadavatel, aby voicebot sám vydával nebo přijímal hlasové příkazy vůči jiným systémům - např. „Zadej objednávku“, „Pošli fakturu“?*

- o Kterých částí série ISO/IEC 30122 se požadavek konkrétně týká a jaká forma prokázání souladu je akceptovatelná?*
- o Je přípustné doložit ekvivalentní metodiky/design guidelines, pokud řešení využívá LLM-řízený dialog bez pevně definovaných hlasových příkazů?“*

Znění vysvětlení zadávací dokumentace č.2:

Požadavek na soulad s ISO/IEC 30122 je uveden pro zajištění kompatibility s běžnými standardy v oblasti hlasových systémů, zejména z hlediska struktury hlasových interakcí, bezpečnosti a ergonomie komunikace.

Cílem není, aby VoiceBot aktivně vydával nebo přijímal přímé hlasové příkazy vůči dalším systémům (např. „Pošli fakturu“), ale aby byl jeho návrh a logika dialogu v souladu s principy standardu – tj. aby rozhraní bylo konzistentní, srozumitelné a bezpečné pro uživatele.

Za dostačující se považuje ekvivalentní metodika či design guidelines prokazující, že řešení dodržuje zásady přehledné hlasové komunikace i v případě využití modelů řízených umělou inteligencí (LLM-based dialog).

Znění žádosti vysvětlení zadávací dokumentace č.3:

„3. Latence

Požadavek na max. 400 ms odezva mezi vstupem a výstupem -- (Příl. č. 4a, bod 3 Technické požadavky). Prosíme o upřesnění:

- o Je zadavatel otevřen úpravě hodnoty tohoto parametru, vzhledem k tomu, že v odvětví je za standardní považována latence okolo 2 000 ms, zejména u řešení využívajících externí API či cloudové služby?*
- o Je zadavatel otevřen úpravě hodnoty (např. vyšší latence při CRM integraci), případně diferenciaci KPI dle scénáře?“*

Znění vysvětlení zadávací dokumentace č.3:

Hodnota latence do 400 ms vychází z požadavku na reálnou plynulost konverzace a okamžitou reakci systému, která je zásadní pro přirozený dialog s uživatelem. Latence v řádu 2 000 ms (2 s) je z pohledu uživatelského zážitku již vnímatelná jako nepřirozené zpoždění a pro cílový charakter řešení (real-time VoiceBot s prodejním zaměřením) není akceptovatelná.

Zadavatel má ambici dosáhnout parametrů srovnatelných se špičkovými systémy, jako jsou Deepgram Aura (<300 ms), AssemblyAI Realtime API (~250–400 ms) nebo NVIDIA Riva (cca 200 ms on-prem). Tyto hodnoty jsou běžně dosažitelné při optimalizaci pipeline (ASR-NLP-TTS) a vhodné architektuře, včetně využití edge computing nebo hybridního nasazení.

Přípustná je dílčí variabilita dle scénáře – např. mírně vyšší latence při hlubších CRM integracích – avšak cílový návrh systému má směřovat k reálné interakci s odezvou v řádu stovek milisekund, nikoli sekund.

Znění žádosti vysvětlení zadávací dokumentace č.4:

„4. Integrace

Požadavek na otevřená REST API a webhooks pro napojení na CRM/databáze, např. Salesforce, HubSpot, ... MySQL, MSSQL, PostgreSQL -- (Příl. č. 4a, bod 3 Technické požadavky). Prosíme o upřesnění:

- o Jaký je účel tohoto požadavku?*
- o [název společnosti] poskytuje end-to-end produkt, tj. kompletní řešení včetně rozpoznávání řeči, orchestrace a telefonní infrastruktury. Naše API jsou určena pouze tam, kde to dává funkční smysl (např. pro vytočení hovoru nebo předání výsledku). Může zadavatel potvrdit, že požadavek na „otevřenou API“ se týká pouze těchto integračních bodů?*

o Může zadavatel specifikovat, které konkrétní integrační toky nebo datové výměny tímto požadavkem zamýšlí?“

Znění vysvětlení zadávací dokumentace č.4:

Požadavek na otevřená REST API a webhooky vychází z potřeby modulární architektury, která umožní VoiceBota napojit na různé CRM systémy a interní datové zdroje, bez závislosti na konkrétním vendorovi. Cílem je, aby řešení umožňovalo flexibilní integraci – např. pro synchronizaci zákaznických dat, předávání výsledků hovorů nebo spouštění obchodních procesů.

Zadavatel potvrzuje, že požadavek na „otevřená API“ se týká zejména integračních bodů – tedy vstupu a výstupu dat, přenosu výsledků (např. transcript, sentiment, CRM ID apod.) a možnosti bezpečně komunikovat s nadřazenými systémy.

Požadavek nesměruje k zásahu do vnitřní architektury dodavatele, ale k zajištění otevřenosti řešení pro další rozvoj a rozšíření.

Příkladem cílových integračních toků je např. přenos metadat o hovoru do CRM (Salesforce, HubSpot), zápis výsledku analýzy do databáze (MySQL, PostgreSQL) nebo notifikace o ukončení hovoru přes webhook.

Znění žádosti vysvětlení zadávací dokumentace č.5:

„5. Vlastnictví zdrojového kódu

Požadavek na i) plné vlastnictví zdrojového kódu -- (Příl. č. 4a, bod 3 a bod 9 Předávací výstupy) a ii) „postoupení veškerých majetkových práv... předání zdrojového kódu.. -- Příl. č. 1a, čl. VII – Duševní vlastnictví). Prosíme o upřesnění:

o [název společnosti] nabízí hotové řešení – platformu [...], kterou pro Gitano cosmetics upravíme a napojíme podle konkrétních požadavků projektu. Na této platformě bude navržena a provozována konverzační logika voicebota. Zdrojový kód samotné platformy zůstává naším vlastnictvím, protože jde o produktovou část systému poskytovanou formou licence, nikoli o vývoj na míru.

o Může zadavatel potvrdit, že požadavek na předání zdrojových kódů se vztahuje pouze na business flows v platformě [název platformy], komponenty, úpravy a integrace vyvinuté specificky pro Gitano cosmetics?“

Znění vysvětlení zadávací dokumentace č.5:

Požadavek na vlastnictví zdrojového kódu je zásadní z hlediska dlouhodobé udržitelnosti, bezpečnosti a škálovatelnosti projektu. Zadavatel potřebuje mít plnou kontrolu nad všemi částmi systému, které byly vyvinuty specificky pro tento projekt – zejména nad konverzační logikou, integračními moduly, orchestrace layerem a nad algoritmy přizpůsobenými interním procesům.

Zadavatel chápe, že generická platforma dodavatele může být licencovaným produktem. Nicméně vše, co bude vyvinuto, upraveno nebo přizpůsobeno na míru pro Gitano Cosmetics, musí být předáno včetně zdrojového kódu a práv k jeho dalšímu rozvoji.

Tento přístup je nutný, aby mohl být VoiceBot dále rozvíjen, škálován a případně provozován jako vlastní SaaS řešení pod správou Gitano. A je také podmínkou dotace, aby příjemce vlastnil zdrojový kód.

Znění žádosti vysvětlení zadávací dokumentace č.6:

„6. Přesnost ASR – min. 90% (CZ/EN)

Požadavek na min. 90 % přesnost v češtině i angličtině (ověřeno na 100 testovacích hovorech poskytnutých zadavatelem) -- (Příl. č. 4a, bod 4).Prosíme o upřesnění:

o Tuto přesnost nelze garantovat, protože jde o parametr poskytovatele služby rozpoznávání řeči (Speech-to-Text). [název společnosti] využívá komerční STT technologie třetích stran, jejichž přesnost se může lišit v závislosti na akcentu, kvalitě zvuku nebo typu hovoru. Může zadavatel potvrdit, že tato hodnota je orientační a nebude předmětem smluvního KPI?

o Je záměrem zadavatele hodnotit úspěšnost projektu podle přesnosti rozpoznávání (ASR), nebo spíše podle celkové úspěšnosti voicebota při porozumění záměru a dosažení výsledku hovoru, což je klíčové z hlediska kvality služby?

o Pokud bude nutné stanovovat KPI na úrovni ASR, prosíme o doplnění metodiky měření a jasné definice testovacího datasetu.“

Znění vysvětlení zadávací dokumentace č.6:

Požadavek na přesnost rozpoznávání řeči (ASR) min. 90 % v češtině i angličtině je stanoven jako minimální funkční cíl pro dosažení přirozené a spolehlivé konverzace. Zadavatel si je vědom, že přesnost může být ovlivněna různými faktory (např. akcent, kvalita signálu), ale očekává, že systém bude optimalizován tak, aby těchto hodnot dosahoval v běžném provozu.

Hodnota 90 % není orientační — vyjadřuje očekávanou úroveň kvality produkčního řešení. Hodnocení projektu však nebude založeno výhradně na přesnosti ASR, ale také na celkové úspěšnosti komunikace (porozumění záměru a správné reakce VoiceBota).

Pro testování budou využity reálné hovory poskytnuté zadavatelem, které reprezentují typické scénáře, doplněné o testovací sadu pro kontrolu konzistence výsledků.

Znění žádosti vysvětlení zadávací dokumentace č.7:

„7. Úspěšnost konverze – min. 70 % průměru lidského operátora po 3 měsících Požadavek v bodu 4 „Výkonové a kvalitativní parametry“ (Příl. č. 4a) -- Prosíme o upřesnění:

o Tento parametr nelze garantovat, protože úspěšnost konverze zásadně závisí na samotném obchodním procesu, typu kampaně, cílové skupině a kvalitě vstupních dat. Může zadavatel potvrdit, že uvedená hodnota je orientační a nebude předmětem smluvní sankce?

o Jak bude úspěšnost měřena a jaká metodika bude použita – např. podle počtu uzavřených obchodů, přepojení, nebo jiného ukazatele?

o Může být KPI stanoveno až po pilotní fázi, kdy bude možné reálně vyhodnotit chování voicebota a srovnat ho s výkonem lidských operátorů?“

Znění vysvětlení zadávací dokumentace č.7:

Požadavek na úspěšnost konverze na úrovni minimálně 70 % průměrného výkonu lidského operátora po třech měsících provozu je definován jako cílová hodnota pro vyhodnocení obchodní efektivity VoiceBota. Cílem není stanovit sankční KPI, ale motivovat dodavatele k optimalizaci interakce, jazykových modelů a scénářů tak, aby VoiceBot reálně dosahoval srovnatelných výsledků s člověkem v prodejním kontextu.

Měření bude probíhat podle reálně dosažených výsledků hovorů – např. míra dokončených prodejků, kvalifikovaných leadů, či úspěšného přepojení na obchodníka.

Po pilotní fázi (cca 3 měsíce provozu) budou data společně vyhodnocena a na jejich základě může být KPI upřesněno nebo upraveno, aby reflektovalo konkrétní podmínky kampaní a chování zákazníků.

Znění žádosti vysvětlení zadávací dokumentace č.8:

„8. Continuous Learning / RL / Transfer / Active Learning

Požadavky v bodu 7 „Mechanismy učení“ (Příl. č. 4a). (Příl. č. 4a) -- Prosíme o upřesnění:

o Uvedený požadavek aktuálně není součástí našeho řešení a ani v tomto roce nebude dostupný. Prosíme o upřesnění, zda se jedná o povinný požadavek pro akceptaci nabídky, nebo o volitelnou (rozvojovou) funkcionalitu do budoucna.“

Znění vysvětlení zadávací dokumentace č.8:

Mechanismy průběžného učení (Continuous Learning, Reinforcement Learning, Transfer Learning, Active Learning) jsou povinnou součástí cílového řešení.

Tyto funkce jsou klíčové pro dosažení dlouhodobé adaptability systému a jeho schopnosti zlepšovat výkon na základě reálných interakcí.

Zadavatel požaduje, aby architektura VoiceBota umožňovala automatické učení z nových dat, vyhodnocování úspěšnosti odpovědí a postupnou optimalizaci modelů – např. pomocí řízeného retrainingu, analýzy chyb a zpětné vazby z reálných konverzací.

Zadavatel je si vědom, že rozsah implementace může být rozfázován (např. základní RL smyčka v 1. fázi, aktivní učení a transfer learning v 2. fázi), avšak princip self-learning architektury je pro přijetí nabídky klíčový.

Cílem je, aby VoiceBot nebyl statický, ale schopný se dlouhodobě zlepšovat a přizpůsobovat trhu a zákazníkům.

„Zvýšení digitální úrovně ve společnosti Gitano cosmetics a.s. – část 2 – Serverová infrastruktura“

Zadavatel poskytuje toto následující vysvětlení zadávací dokumentace vztahujícím se k výše uvedené zakázce.

Znění žádosti vysvětlení zadávací dokumentace č.1:

„9. *Rozsah a účel řešení*

Požadavky v Příloze č. 4b – Specifikace předmětu plnění – část 2“ stanovují parametry, které odpovídají vysoce výkonným HPC systémům. -- Prosíme o upřesnění:

o Jaký je účel požadované infrastruktury – má být primárně určena pro provoz AI voicebota (část 1) nebo má sloužit i pro jiné interní systémy společnosti Gitano cosmetics?

o Potvrzuje zadavatel, že řešení má být nasazeno on-premise v místě provozovny v Dubí a nebude možné nabídnout alternativní cloudovou nebo hybridní architekturu?

o Může zadavatel upřesnit, zda požaduje reálné fyzické dodání hardwaru, nebo postačí pronájem / výpočetní kapacita jako služba (IaaS)? o Jaká je minimální očekávaná zátěž systému – zejména počet současných hovorů, uživatelů nebo AI úloh, ze kterých tyto výkonnostní parametry vychází?

10. *Návrh k přehodnocení požadovaných parametrů*

Navrhujeme, aby zadavatel přehodnotil navrhované parametry a zvážil jejich úpravu tak, aby odpovídaly reálným možnostem současných komerčně dostupných řešení:

- *výpočetní uzly se 4–8 GPU,*
- *operační paměť v rozsahu 256–512 GB RAM na uzel,*
- *síťovou propustnost 25–100 Gb/s,*
- *kapacitu pro stovky až nižší tisíce paralelních hlasových hovorů, - firewall s průchodností 10–40 Gb/s.“*

Znění vysvětlení zadávací dokumentace č.1:

Odůvodnění technických parametrů části : Serverová infrastruktura

Naše společnost potvrzuje, že cílem projektu je vytvořit škálovatelnou AI VoiceBot platformu, která bude nejen interně využívána, ale také komerčně poskytována jako SaaS řešení dalším subjektům na mezinárodní úrovni. Z tohoto důvodu je navržená architektura serverové infrastruktury koncipována jako výkonný a dlouhodobě udržitelný základ pro rozvoj vlastního AI jádra, včetně trénování a provozu modelů v reálném čase.

Hlavní důvody, proč požadované parametry odpovídají high-performance computing (HPC) úrovni:

Reálný čas a nízká latence (ultra-low latency)

VoiceBot musí zpracovávat a generovat řeč v reálném čase s latencí pod 400 ms při současném běhu stovek až tisíců hovorů. To vyžaduje výpočetní výkon GPU clusteru, který překračuje standardní parametry běžných komerčních serverů.

Multimodální a emocionální analýza

Platforma zahrnuje rozpoznávání emocí, analýzu tónu hlasu, real-time diarizaci a adaptivní modelování hlasu. Tyto úlohy kombinují STT, TTS a NLP inference vrstvy, které jsou výrazně výpočetně náročné a GPU-intenzivní. Trénování a fine-tuning modelů

Na rozdíl od běžného provozu SaaS chatbotů zde probíhá kontinuální trénování vlastních jazykových a akustických modelů na interních datech, což vyžaduje vyšší počet GPU a rozsáhlou RAM. Cílem je nezávislost na externích poskytovatelích a možnost lokálního trénování modelů.

Škálovatelnost a internacionalizace

Systém je koncipován tak, aby podporoval různé jazykové mutace (EN, CZ později DE, FR...) a mohl být distribuován formou white-label SaaS pro zahraniční klienty. Předpokládáme postupné rozšíření na tisíce současně aktivních hovorů v horizontu 3–5 let, což vyžaduje robustní architekturu už v počáteční fázi.

Bezpečnost a kontrola dat (on-premise deployment)

Pro fázi vývoje a testování je nezbytné, aby řešení běželo on-premise v Dubí, kvůli požadavkům GDPR a kontrole citlivých dat z reálných hovorů. Cloudové řešení v této fázi není akceptovatelné, i když do budoucna bude možné zvažovat hybridní architekturu.

Pro ilustraci uvádíme, že naše stávající callcentrum zvládá v současnosti přibližně 800 paralelních hovorů. V okamžiku, kdy bude VoiceBot nasazen nejen interně, ale také u dalších klientů (např. jiných callcenter nebo partnerských firem), bude tento počet násoben mnohonásobně – což vyžaduje odpovídající výpočetní kapacitu již v základní infrastruktuře.

Z výše uvedených důvodů není vhodné snižovat požadované parametry. Současné hodnoty (4–8 GPU na uzel, RAM 256–512 GB atd.) představují minimální základ pro efektivní trénování a real-time inference více paralelních modelů. Jakákoli redukce by zásadně omezila schopnost systému zvládat dynamické škálování a snížila by jeho konkurenceschopnost na mezinárodním trhu.

Informace o prodloužení lhůty pro podání nabídek:

Zadavatel prodlužuje lhůtu pro podání nabídek **do 19. 11. 2025**

Otevírání elektronických nabídek proběhne dne **19. 11. 2025 v 10:05 hodin**.

V Brně, dne 5. 11. 2025

Deregio Tender, s.r.o.

Ing. Jan Ševčík, jednatel

Zástupce zadavatele