Příloha č. 1 Zadání pro zpracování analýzy možností implementace pokročilých nástrojů (včetně AI) do portálů CzT

# Úvod & Business Summary

Tento dokument obsahuje popis problémů, se kterými se v běžném provozu setkává redakce portálů Kudy z nudy a #VisitCzechia, nastiňuje jejich řešení za předpokladu využití již existujících dat a jejich následné zpracování umělou inteligencí nebo jinými pokročilými nástroji. Objednatel předpokládá, že v rámci Analýzy bude tento dokument detailně rozpracován a budou navržena opatření, která pomohou k redukci resp. celkové eliminaci těchto problémů. Objednatel dále předpokládá, že Zhotovitel posoudí proveditelnost řešení, jeho efektivitu, soulad s příslušnými zákony a pravidly webové etiky.

Požadované výstupy Analýzy budou v souladu s Přílohou č. 3 Smlouvy Nabídka Born Digital a zároveň jejich minimální rozsah bude následující:

* návrh technického řešení formou technického zadání za využití již existujících pokročilých nástrojů běžně dostupných na trhu nebo formou prototypu v nástrojích na míru,
* studie proveditelnosti, přičemž jedním z aspektů je realizovatelnost technického řešení v prostředí Kentico CMS pomocí externích API apod.,
* odhad chybovosti výstupů generovaných pokročilými nástroji,
* odhad nákladů na vývoj a implementaci samotného usecase a odhad nákladů na provoz se zvážením možných alternativ,
* odhad úspor včetně výpočtu návratnosti investovaných prostředků (včetně nákladů na tuto analýzu),
* analýza právních aspektů zahrnující popis rizik v souvislosti s implementací navrhovaného řešení a návrh opatření vedoucích k jejich eliminaci,
* analýza navrhovaného řešení z pohledu pravidel webové etiky.

Požadované výstupy Analýzy budou využity jako podklad pro rozhodnutí Objednatele, zda, jakým způsobem, za jakých podmínek a kdy provede implementaci pokročilých nástrojů (včetně AI) do portálů Objednatele s cílem zefektivnit provoz portálů Objednatele.

# Formulace problému redakcí portálů

|  |  |  |
| --- | --- | --- |
| **ID** | **Název** | **Popis** |
| P–1 | Nedostatečné personální kapacity na schvalování obsahu | Každý nově vložený článek na Kudy z nudy prochází manuálním schvalováním redakce. Denně je nově zadáno zhruba velké množství záznamů, přičemž redakce není. v tuto chvíli schopna toto množství plně odbavit.  |
| P-2 | Pravopisné chyby a rozdílná stylistika | Na Kudy z nudy přispívá velké množství externích autorů, kteří mají různou úroveň českého jazyka (pravopis a stylistika textu) a marketingové komunikace. |
| P-3 | Subjekty inzerované na portálech CZT neplní publikační podmínky | Publikační podmínky jsou umístěny na <https://www.kudyznudy.cz/faq-casto-kladene-otazky/registrace-a-nove-zaznamy>. Nejčastějšími problémy jsou přidávání plakátů k záznamům, používání velkých písmen a neumisťování banneru či loga s odkazem www.kudyznudy.cz na webových stránkách provozovatele. |
| P-4 | Vkládaná data obsahují chyby | Jedná se například o chyby pravopisné a dále chyby „z nepozornosti“, kdy je např. zadána jen část tel. čísla atd..  |
| P-5 | Zatraktivnění obsahu pro návštěvníka portálu | Atraktivita portálu je dána kvalitou obsahu a použitými formami. Kromě interních dat je vhodné využívat data třetích stran. CzT má zájem na tom neustále hledat nové možnosti pro zatraktivnění portálů prostřednictvím validních dat třetích stran.  |
| P-6 | SEO webových portálů | Jeden z hlavních cílů CzT je visibilita portálů. SEO je kontinuální proces, který spotřebovává velkou část kapacit. CzT neustále hledá řešení, která pomáhají tento proces maximálně zefektivnit.  |
| P-7 | Zlepšení UX rozhraní pro vkládání záznamů (akce a aktivity) na KZN  | Současný formulář je již zastaralý a vkladatelům nemusí vyhovovat. Pokud se jim nedaří záznam vložit, zatěžují svými dotazy redakci. Zároveň Objednatel plánuje vytvoření centrálního datového skladu, nad kterým bude rozhraní nově vytvořeno.  |

# Představa řešení problému

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Název** | **Problém** | **Priorita** | **Popis** |
| Ř–1 | Automatická validace | P-1 | 1 | Údaje o jednotlivých subjektech na portále CzechToursim (Kudy z nudy, VisitCzechia) jsou dohledatelné a porovnatelné s údaji stejných subjektů umístěných na jiných portálech (například adresa provozovatele, otevírací doba aj.). Cílem je, aby data subjektu na našich portálech byla automaticky validována za pomoci nejvhodnějšího z externích zdrojů. |
| Ř-2 | Automatická kontrola pravopisu a sjednocení stylistiky | P-2 | 2 | Cílem je implementace nástrojů, které odhalí a opraví gramatické chyby, překlepy a upraví článek napsaný autorem tak aby svým stylem odpovídal co nejvíce stylu psaní na webovém portálu. |
| Ř-3 | Vhodnost akce k publikaci na webovém portálu | P-1 | 3 | Na Kudy z nudy se hromadí spousta irelevantních nabídek, například nabídka finančních služeb, členství v sektách a dalších produktů nesouvisející s cestovním ruchem. Cílem je tyto akce nesouviseních s cestovním ruchem nezveřejnit na Kudy z nudy. |
| Ř-4 | Automatická kontrola publikačních podmínek | P-1, P-3 | 1 | CzechTourism má nastavené podmínky pro publikaci obsahu na webových portálech. Jednou z nich je povinnost subjektu umístit na své webové stránce logo s proklikem na příslušný webový portál CzechTourism. Cílem je automaticky kontrolovat splnění těchto publikačních podmínek (včetně správného využití brandmanuálu) a automaticky notifikovat subjekt a redakci o nesplnění těchto podmínek.  |
| Ř-5 | Automatický rating subjektů | P-1, P-2, P-4 | 2 | Cílem je, aby vytvořit systém (nástroj), který zaznamenává chybovost (chybná technická data, pravopisné chyby atd.) jednotlivých subjektů a na základě ní vypočítává skóre subjektu. Zadané články subjektu s vysokým skóre nemusí být tak často redakcí kontrolovány jak manuálně, tak automaticky (úspora kapacit tak výpočetního výkonu).  |
| Ř-6 | Automatická aktualizace akcí | P-1 | 1 | Většina akcí vložených na portálech CzechTourism se opakuje (např. každoročně). Cílem je po skončení akce detekovat, zda se bude konat i budoucnu (z externích zdrojů) a aktualizovat akci jak po stránce obsahové (program aj.) tak po stránce technické (datum konání, kontakt na pořadatele apod.) a vytvořit koncept této nové akce.  |
| Ř-7 | Vyhledávání nových akcí | P-1 | 2 | Cílem je z externích zdrojů najít akce, které zatím nejsou v katalogu Kudy z nudy nebo VisitCzechia a z externích zdrojů vytvořit koncept této akce.  |
| Ř-8 | Kopírování akcí mezi VisitCzechia a Kudy z nudy | P-1 | 1 | V rámci redakce webových portálů jsou dvě oddělené redakce pro Kudy z nudy a VisitCzechia. Cílem je „vytěžit“ obsah Kudy z nudy pro potřeby VisitCzechia. Vzhledem k tomu, že obsah článků na VisitCzechia je s přihlédnutím na rozdílné cílové skupiny odlišný (zahraniční vs. čeští cestovatelé), je potřeba články převzaté z Kudy z nudy vhodně přeformulovat za využití AI a externích zdrojů dat.  |
| Ř-9 | Propojování externích zdrojů dat | P-5 | 1 | Cílem je propojení a import dostupných externích dat (např. hodnocení) z Google Places, Booking apod. |
| Ř-10 | Duplicita obsahu, včetně fotek | P-1 | 2 | V rámci Kudy z nudy využíváme importy třetích stran, kdy je takto vložený obsah mnohdy duplicitní. Cílem je tedy vytvoření systému, který rozpozná, že tento obsah již v rámci webového portálu existuje a redaktora na toto upozorní. V rámci VisitCzechia naopak existuje velké množství redaktorů, kteří vkládají fotografie k jednotlivým záznamům a v Media Library CMS systému jsou vloženy duplicitně. Cílem tedy je rozpoznat, že tato fotografie je již v Media Library CMS uložena. Zároveň by bylo vhodné nově vloženou fotografii automaticky pojmenovávat na základě dat z interní Digital Media Library (media.visitczechia.com) a Shutterstock. |
| Ř-11 | Štítkování obsahu | P-6 | 2 | Cílem je vytvoření systému štítků pro web VisitCzechia. Štítky by měly být automaticky generované v závislosti na obsahu daného článku s možností zásahu redaktora. Pro každý štítek by měla vzniknout landing page s odkazy na stránky, kde je tento štítek použit. Vygenerovaná klíčová slova formou štítků by měla být zahrnuta v meta keywords. Bude nezbytná konzultace se specialistou na SEO. |
| Ř-13 | Optimalizace výkonových kampaní a SEO | P-6 | 3 | Algoritmy Google v současné době procházejí neustálým vývojem s čím dál větší mírou zapojení umělé inteligence. Cílem je nasazení umělé inteligence, která dokáže reagovat na tyto změny v oblasti Google vyhledávání a placených kampaní. Například Google zapojuje do popisu ve vyhledávání i description z jiných interně pro linkovaných stránek. |
| Ř-14 | Automatické vyhledávání fotek k akcím / aktivitám | P-1 | 3 | V současné době nestačí kapacity redakce k doplňování nových fotek z externích fotobank k současným záznamům (akcím a aktivitám) na Kudy z nudy. Cílem je automatizovaným způsobem tyto externí fotobanky (Shutterstock, Profi Media) procházet a vytipovávat fotky, které by bylo vhodné k článkům doplnit.  |
| Ř-15 | Interní prolinkování | P-1, P-6 | 2 | V současné době je na portálech velké množství obsahu, který je vhodné pro zlepšování SEO propojovat do obsahů nově aktualizovaných nebo přidaných. Cílem je tedy vytvoření systému, který v novém obsahu vytipuje pasáže, které by se daly vhodně pro linkovat na již jiný existující obsah v rámci obou webových portálů. |
| Ř-16 | Upgrade rozhraní pro vkládání akcí a aktivit do KZN  | P-7 | 1 | Současné rozhraní je potřeba nově navrhnout tak, aby editor co nejrychleji prošel procesem vložení záznamu. Je potřeba, aby jej systém inteligentně naváděl a nebyl zatěžován irelevantními dotazy resp. aby jej opravoval v případě nesprávného postupu.  |