**PŘÍLOHA Č. 8 – ZÁVAZNÉ TECHNICKÉ PARAMETRY PŘEDMĚTU VEŘEJNÉ ZAKÁZKY**

V následující tabulce je uveden výčet závazných parametrů poptávaného řešení včetně jejich popisu. Bezvýhradní naplnění těchto parametrů je nutnou podmínkou pro další hodnocení nabídek. Nabídky, které prokazatelným způsobem nenaplní závazné parametry, nebudou hodnoceny a budou ze zadávacího řízení vyřazeny.

|  |  |  |
| --- | --- | --- |
| **Název parametru** | **Popis požadované funkcionality** | **Detailní popis způsobu naplnění** |
| Definice nových zdrojů dat | Poskytovatel služby umožní vložení a definici nového zdroje dat, jehož extrakce bude následně součástí poskytovaných datových výstupů |  |
| Požadované výstupy služby | Popis způsobu zajištění výstupu extrakčních dat ve formátech XML, CSV, JSON |  |
| Extrakce http i https stránek | Extrakce dat musí umožňovat práci se zabezpečeným protokolem HTTPS |  |
| Extrakce nezaheslovaných zip, rar archivů | Popis a způsob zajištění extrakce informací z nezaheslovaných archivačních souborů typu zip, rar obsahující dokumenty typu PDF, MS Excel, MS Word, MS PPT |  |
| Extrakce textu z dokumentů PDF, MS Excel, MS Word, MS PPT | Popis a způsob zajištění extrakce informací z dokumentů formátu PDF, PPT, MS Word, MS Excel a to včetně extrakce metadat |  |
| Podporované formáty zpracovávaných dokumentů | Podporované formáty: XML, PDF, MS Office, Open Office, OpenXML |  |
| Vstupní rozhraní | Vstup dat ze souborového systému a pomocí REST CRUD, JSON a HTTP API s možností definice parametrů v rámci administrace aplikace. |  |
| Nastavitelné přístupové práva na základě zdroje extrahovaných dat | Popsat návrh a technologii řešení přístupových práv k extrahovaným informacím |  |

|  |  |  |
| --- | --- | --- |
| Výstupní rozhraní | Návrh definice výstupního rozhraní pro analýzy uložených dat. Požadovány jsou formáty REST CRUD (XML), JSON, HTTP |  |
| Pokročilé vyhledávání | Pokročilá schopnost práce se seznamy vyloučených slov a synonym |  |
| Musí nabízet zpracování vstupních dat do formy vhodné pro analýzu dat | Čištění a normalizace vstupních dat |  |
| Extrakce metadat |  |
| Lingvistické zpracování textů v českém jazyce, angličtině a němčině |  |
| Plná podpora zpracování přirozeného jazyka (NLP) v českém jazyce |  |
| Skloňování a časování (lematizace) českých slov, ne pouze redukce na kořen slov (stemming) |  |
| Rozpoznávání vět a větných členů |  |
| Tvorba a správa kategorií a pravidel pro automatické zařazování obsahu | Uživatelsky přívětivý, grafický nástroj pro vytváření a změny pravidel kategorizace v českém jazyce bez nutnosti dodatečného programování |  |
| Tvorba uživatelských pravidel |  |
| Tvorba znalostních kategorií výčtem klíčových slov |  |
| Omezení výběru dle definovaných kritérií | V rámci prezentační vrstvy požadujeme možnost třídit dle metadat zpráv (například autor), datum pořízení, zdroje nebo automaticky generovaného konceptu zpráv |  |
| Exporty sestav | Aplikace umožní export zdrojových dat do běžně dostupných kancelářských programů jako je MS Office Excel, MS Office Word |  |
| Podporované vstupní formáty aplikace | Aplikace podporuje a zpracovává následující vstupní datové formáty: XML, JSON |  |

|  |  |  |
| --- | --- | --- |
| Sumarizace témat s možností prokliku na detail | Zachycení vrcholové informace s možností prokliku na detail zpráv až na úroveň samotných vstupů dat |  |
| Automatické generování kontextu a souvisejících dokumentů | Aplikace umožní na základě analyzovaného tématu dovodit související dokumenty a zobrazit je prostřednictvím přehledové informace |  |
| Podpora běhu aplikace v rámci poskytování služby | Zajištění technické podpory pro běhu služby v pracovních dnech od 8 do 18 hodin po dobu délky poskytování této služby |  |
| Dostupnost služby | Definice a přesný popis způsobu zajištění požadované dostupnosti poskytované služby 24x7x365 |  |
| Sada grafů umožňující grafickou interpretaci trendů | Popis a detailní návrh nejvhodnějších typů grafických prostředků umožňující zachycení trendových veličin |  |
| Zajištění zabezpečeného provozu služby | Poskytování služby musí být realizováno prostřednictvím https protokolu s platným certifikátem |  |
| Dostupnost poskytované služby | Požadovaná dostupnost služby je v režimu 24x7x365 (24 hodin, 7 dní v týdnu, 365 dní v roce) |  |
| Extrakce dat z prostředí sociálních sítí facebook, twitter, youtube a dalších | Popis způsobu a technologického řešení umožňující nejlepší možný způsob extrakce dat a informací z prostředí požadovaných sociálních sítí, včetně způsobů řešení specifických problémů, které tyto zdroje dat mají. |  |
| Způsob zajištění technické podpory pro zajištění běhu služby | Detailní popis způsobů zajištění technické podpory pro zajištění běhu služby v požadovaném čase každého pracovního dne mezi 8-18:00 |  |