

Příloha č.5

ČÁST V. - DATOVÉ SKLADY, MANAŽERSKÉ INFORMAČNÍ SYSTÉMY A NÁSTROJE BUSINESS INTELLIGENCE

Verze dokumentu

Následující tabulka obsahuje informace o verzi dokumentu:

Verze	Datum	Popis změny v dokumentu
0.1	18. 7. 2010	První draft
1.0	30. 7. 2010	Druhý draft
1.2	3. 8. 2010	Dokument k připomínkování -technické kapitoly
1.4	8. 8. 2010	Doplnění obecných kapitol
1.5	10. 8. 2010	Doplnění nových kapitol
2.1	15. 8. 2010	Doplnění nových kapitol
2.2	16. 8. 2010	Formální úpravy
2.4	17.8.2010	Příprava dokumentu k připomínkování
3.1	18.8.2010	Kompletní doplnění kapitol
4.0	20.8.2010	Změna struktury dokumentu
5.0	25.8.2010	Odstranění kapitol v souladu se snížením rozpočtu
5.1	28.8.2010	Formální úpravy, verze k akceptaci
5.2	31.8.2010	Variantní řešení
6.0	4.9.2010	Zpracování změn z 3.9.2010
6.1	9.9.2010	Úprava hodnocení variant
6.2	12.9.2010	Zpracování připomínek, úprava seznamu dat.tržišť
6.3	16.9.2010	Úprava seznamu dat.tržišť

Tabulka 1: Verze dokumentu

Obsah

1	Technické řešení.....	5
1.1	Vlastní koncepce řešení	5
1.1.1	Návrh a popis architektury řešení	5
1.1.2	Variantské návrhy technického řešení – HW/SW/data.....	16
1.1.3	Naplnění požadavků typizovaného projektu.....	17
1.2	Porovnání variant technologických řešení	17
1.2.1	Srovnání nabídek jednotlivých dodavatelů (poptávkové řízení).....	17
1.2.2	Výhody a nevýhody jednotlivých řešení.....	18
1.2.3	Analýza technických a bezpečnostních rizik	18
1.3	Doporučení a upřesnění pro účely zadávací dokumentace a realizační projektové dokumentace.....	19
1.3.1	Specifikace zadání technického řešení	19
1.3.2	Požadavky na implementaci, školení a technickou podporu.....	19
1.4	Provozní zajištění projektu	20
1.4.1	Potřebné energetické a materiálové toky.....	20
1.4.2	Záruky a servis	20
1.4.3	Údržba a nákladnost oprav	20
1.4.4	Údaje o provozním zajištění SW a datových komponent	20
1.4.5	Změny v provozní náročnosti vlivem opotřebení.....	21

Seznam tabulek

Tabulka 1:	Verze dokumentu.....	2
Tabulka 2:	Metadata datového skladu	11
Tabulka 3:	Metadata úřadu.....	11
Tabulka 4:	Souhrnné databáze ČSU	12
Tabulka 5:	Specifické publikace ČSU	12
Tabulka 6:	Regionální databáze.....	12
Tabulka 7:	Data Eurostatu, OECD a další mezinárodní data	12
Tabulka 8:	Data ČNB.....	12
Tabulka 9:	MPSV	12
Tabulka 10:	Průzkum informační gramotnosti	13

Tabulka 11: Obraty.....	13
Tabulka 12: Doklady.....	13
Tabulka 13: Kapitoly, Příjmy.....	13
Tabulka 14: Cash Flow.....	13
Tabulka 15: Účetnictví.....	13
Tabulka 16: Majetek.....	14
Tabulka 17: Doprava.....	14
Tabulka 18: Registr územní struktury.....	14
Tabulka 19: Registr subjektů.....	14
Tabulka 20: Registr příspěvkových organizací kraje.....	14
Tabulka 21: Registr organizační struktury úřadu.....	15

Seznam obrázků

Obrázek 1: Architektura DWH a BI.....	6
---------------------------------------	---

1 Technické řešení

1.1 Vlastní koncepce řešení

Vlastní koncept řešení vychází z dokumentu „Datové sklady a nástroje Business Intelligence - Typizovaný projektový záměr“, který je přílohou Příručky pro žadatele projektu, ve kterém jsou definované požadavky na řešení i návrhy principu jejich řešení.

1.1.1 Návrh a popis architektury řešení

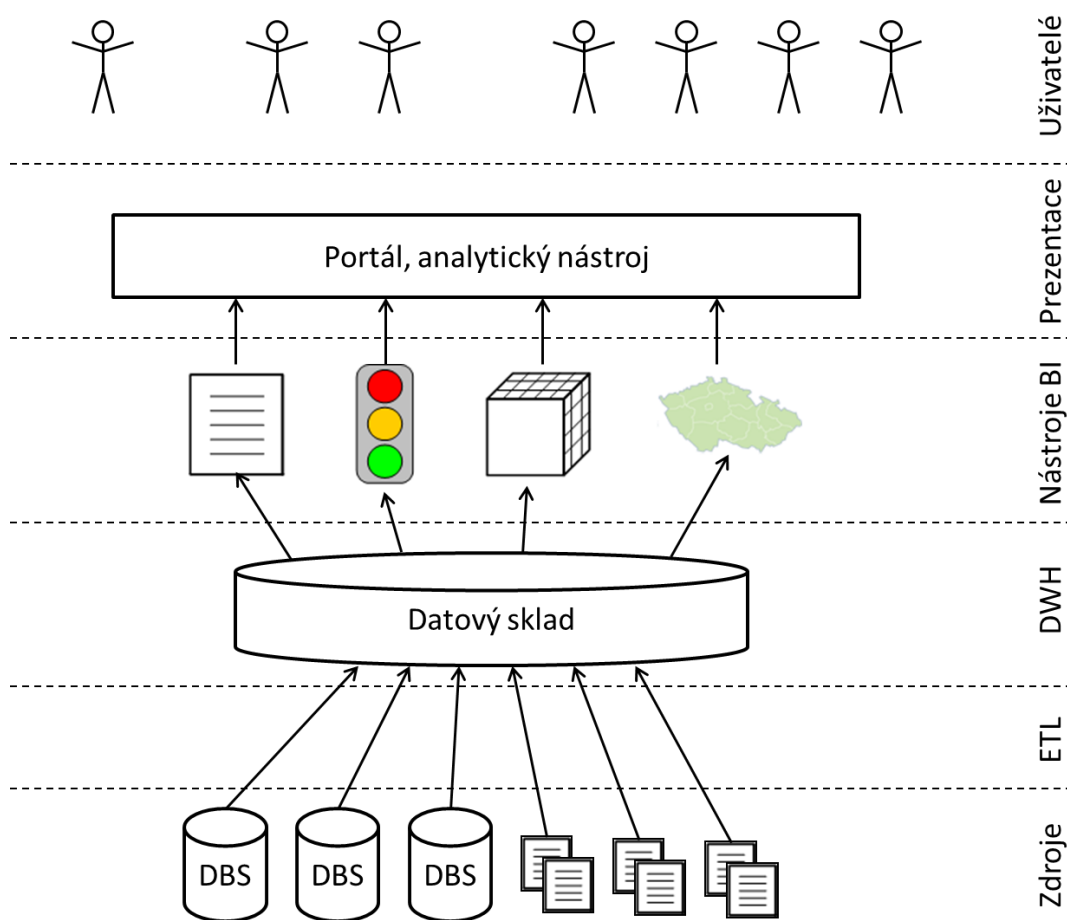
Základní axiomem architektury je rozdělení na tři základní vrstvy:

1. Vrstva transformačních mechanismů: zajištění načítání zdrojových dat ve stanovených intervalech do odděleného datového úložiště
2. Analytická vrstva: transformace načtených dat do podoby vhodné pro rychlou ad-hoc analýzu a reporting, jejich uložení
3. Prezentační vrstva: nasazení analytických, reportovacích, vizualizačních a interaktivních nástrojů pro analýzu, vizualizaci a reportování dat

Podrobný popis architektury je uveden v následujících kapitolách.

1.1.1.1 Datový sklad a nástroje Business Intelligence

V oblasti datového skladu a nástrojů Business Intelligence se v případě KÚ Karlovy Vary jedná o upgrade stávajícího řešení, přičemž je třeba, aby se budoucí dodavatel držel následujících pravidel, které vychází z požadavků typizovaného projektu:



Obrázek 1: Architektura DWH a BI

Ze zdrojových systémů jsou data pomocí ETL procesů pravidelně načítána do fyzicky i logicky odděleného datového úložiště – do datového skladu. Datový model datového skladu je uzpůsoben primárnímu účelu – vyřizování komplexních analytických dotazů. Proto v rámci ETL procesů dochází k čištění, integraci a transformaci načítaných dat. Nad daty uloženými v datovém skladu jsou vytvořeny analytické modely multidimenzionálních datových kostek pro pokročilou ad-hoc analýzu a modely parametrických sestav, které jsou automaticky distribuovány koncovým uživatelům vhodným komunikačním kanálem ve vhodném formátu, který bude určen ve fázi analýzy. Takto postavené řešení poskytuje stabilní základ pro tvorbu dalších, vyšších pater řešení BI – např. Balanced Scorecards, který umožňuje měřit celkovou výkonnost organizace skládáním klíčových výkonnostních ukazatelů do hierarchicky uspořádaných scorecard – pracovník sleduje aktuální stav do té míry agregace, která mu přísluší, vždy ale s možností podrobné ad-hoc analýzy, pokud je třeba.

Klíčovými prvky celého modelu však zůstávají lidé – od pracovníků, kteří pracují s transakčními systémy a vytváří tak zdrojová data, přes správce systému, kteří monitorují všechny prvky řešení, resp. jsou proaktivně informováni o nesrovnalostech a chybách, až po koncové uživatele – analytiku, manažery, lektory, občany – každý z nich má umožněn přístup právě k takovým datům a nástrojům, které v danou chvíli potřebuje.

1.1.1.2 Metodika zajištění základní báze dat, management validního sběru, hodnocení efektivity

1.1.1.3 Zajištění základní báze dat

Metodikou zajištění základní báze dat rozumíme administraci datových tržišť, tzn. jejich údržbu včetně správy různých indikátorů a hierarchických struktur tj. dimenzí a hierarchií. Vůči těmto indikátorům a strukturám (hierarchiím, dimenzím) je vhodné mít zajištěn sběr a kompletnost nutných popisů, nejlépe ve formě strukturované dokumentace (metadat). Je třeba zajistit:

- zabezpečení základního katalogu indikátorů s možností zápisů metodiky, způsobu výpočtu, katalogizace
- zabezpečení základního registru dimenzí a hierarchií datových skladů a datových kostek, tzn. tvorba a údržba vlastních hierarchií a dimenzí,
- správa vazeb mezi různými datovými zdroji na hierarchie, dimenze,
- možnost sledovat libovolné další popisné vlastnosti (atributy) struktur – zabezpečení sběru a kompletnosti dokumentace na základě těchto strukturovaných informací (metadat),
- manuální sběr hodnot do datového skladu (ať už plánované nebo dosažené skutečnosti) je nutné řídit dle hierarchie struktury – je také třeba ovlivnit směr sběru – zda-li top-down i bottom-up (seshora dolů i zespoda nahoru),
- řízení sběru seshora dolů i zespoda nahoru, mít přehled o stavu plánování a rozpracování plánů projektů, cílů atd.
- Ošetřenu vazbu na budoucí legislativní změny s ohledem na probíhající reformu veřejné správy

1.1.1.3.1 Management validního sběru a hodnocení efektivity

Pro účinné hodnocení efektivity je třeba zajistit konfiguraci a administraci kritérií hodnocení, skrze které bude možné řízení sběru plánu, výhledu i skutečnosti a to centrálně pro všechny definované oblasti. Je třeba také disponovat mechanismem, který zajistí validní prokazatelný sběr informací pro oblasti, kde neexistuje validní datový zdroj. K tomu je třeba zajistit:

- administrace jednotlivých oblastí a předmětů hodnocení efektivity (ať už interní služby jako IT, Personalistika atd. tak externí služby jako správa a údržba silnic, základní dopravní obslužnost, sociální služby, školství atd.)
- administrace indikátorů (ukazatelů) hodnocení efektivity, které budou tvořit kritéria hodnocení, jejich zařazení do konkrétních předmětů hodnocení efektivity,

- administrace hodnocených jednotek, které budou hodnoceny v daném předmětu. S tím souvisí i další hierarchie (dimenze), které budou tvořit další rozměr hodnocení efektivity,
- požadavkem je uživatelská přívětivost správy obsahu pro hodnocení efektivity, aby každou oblast mohl mít na starosti i odborný garant, který nemusí mít pokročilejší IT znalosti,
- zabezpečit validní a prokazatelný sběr indikátorů hodnocení efektivity a to jak manuálně tak (polo)automatizovaně např. z účetnictví nebo rozpočtu. Management sběru musí umožnit centrální řízení sběru všech hodnot a to:
 - jak manuálně tak automatem,
 - plánovaných hodnot i dosažené skutečnosti,
 - libovolnou formou např. skrze „plány“, „výkazy“ atp. a to jak online tak offline (např. formulářů technologie 602 XML),
 - sbírané hodnoty musí být garantované
 - sběr bude možné řídit
 - umožnit schválit jednotlivé sebrané hodnoty (nikoliv celé výkazy), nebo je vrátit k opravě, zabezpečit výpočet a analýzu různých druhů odchylek pro vytvoření efektivních controllingových výstupů.

1.1.1.4 Uživatelé a prezentační vrstva

1.1.1.4.1 Uživatelé

V následujících odstavcích jsou uvedeny základní role budoucích uživatelů datového skladu:

Vrcholný management

Vrcholný management očekává od manažerských informačních systému jednoduché, intuitivní ovládání, jasné, jednoduché výstupy agregovaných dat do vyšších celků. Vhodným nástrojem se proto jeví definovat klíčové výkonnostní ukazatele (KPIs), které je dále možné skládat do hierarchických struktur, tzv. Scorecard.

Výstupy pak mohou být vrcholnému managementu prezentovány například formou „semaforů“ nebo jednoduchých tabulek s možností prokliku k detailním údajům.

Výhodou takto postaveného řešení je definice a vazba KPIs na konsolidovaná data použita např. k multidimenzionální analýze. Uživatel má tedy v případě potřeby (ukazatel se například nevyvíjí podle předpokladů) možnost interaktivně přejít k pokročilé OLAP analýze a identifikovat zdroj problému. V praxi je však využití OLAP analýzy vrcholným managementem spíše výjimečné.

Dalším výstupem, který míří k top managementu, jsou automaticky distribuované sestavy.

Střední management

Střední management uvítá veškeré nástroje BI – od výše uvedených Scorecard, kde mohou rychle získat přehled o vývoji důležitých ukazatelů své oblasti (tedy o „úroveň níž“ než top management), přes multidimenzionální OLAP analýzu pro řešení složitějších problémů až po řízený reporting s možností ovlivnění výstupu výběrem parametrů.

Analytik

Analytika lze považovat, ze zkušeností získaných zpětnou vazbou s uživateli podobných řešení, za hlavního vytěžovatele datového skladu, který bude využívat všech možností analytických nástrojů. Je to osoba, která pomáhá ostatním uživatelům pracovat s nástroji BI, vytváří pokročilé ad-hoc analýzy, shromažďuje podklady a připomínky uživatelů, čímž vytváří podklady pro rozhodování o rozvoji řešení Business Intelligence. Zároveň je to tvůrce parametrizovatelných a automaticky distribuovaných sestav, které vznikají na základě požadavků uživatelů.

Analytik může využívat stejně jako kdokoli jiný například analytický portál, ideálním nástrojem však pro něj je pokročilá analytická aplikace, která dokáže naplno využít data uložená v datovém skladu.

Výkonný pracovník

Výkonný pracovník ke své práci potřebuje výstupy, které je možné předem poměrně přesně definovat, stejně tak je možné předvídat i situace, kdy mu má být konkrétní výstup k dispozici. Z těchto důvodů je ideální využít řízeného reportingu.

Systém automaticky zajistí distribuci sestav např. mailem v definovaných časových nebo událostních okamžicích v předem definovaném formátu závislém na skutečné potřebě konkrétní skupiny koncových uživatelů.

Veřejnost

Z pohledu krajského úřadu je informování veřejnosti velmi důležité. Jelikož navržený datový sklad obsahuje mnoho údajů, které jsou veřejné a často jsou veřejností vyžadována, přinese jejich automatická prezentace ulehčení pracovníkům, kteří musí výstupy pro veřejnost připravovat.

Základním přístupovým prostředkem pro veřejnost je webový portál. Maximální důraz je kladen na platformovou nezávislost s minimální mírou podpory koncových uživatelů. Ze zkušenosti jiných krajů jsou primárním nástrojem prezentace dat předem připravené parametrické sestavy zobrazující údaje pomocí tabulek, grafů a map. Zároveň je ale vhodné dát veřejnosti nástroje, které jí přes webové rozhraní umožní vytvořit ad-hoc výstup z multidimenzionálních OLAP struktur např. pomocí jednoduchých průvodců.

1.1.1.4.2 Prezentační vrstva

V následujících odstavcích jsou uvedeny jednotlivé komponenty prezentační vrstvy, resp. jsou uvedeny způsoby využití dat uložených v datovém skladu:

1.1.1.4.2.1 Služby

- Nad datovým skladem bude vytvořeno rozhraní založené na Web Services.
- Pomocí Web Services bude definován katalog statistických a analytických služeb, které bude možné volat z jakéhokoli systému založeného na službách (např. Integrovaná platforma, Workflow nástroj, AIS, apod.)
- Katalog bude publikován také vně technologického centra – i externí organizace (zřizované organizace, ORP) mohou tyto analytické služby volat

1.1.1.4.2.2 Interní přístup

- Reporting
 - přístup k reportům přes webové rozhraní
 - export reportů do různých formátů (xls, pdf, obrázek, text, xml)
 - automatická distribuce reportů
 - pokročilé řízení přístupu uživatelů k reportům i vlastnímu obsahu reportů
 - centrální správa řešení
- Multidimenzionální analýza
 - pokročilá analýza za pomoci klientských nástrojů přístupu (webový portál, speciální aplikace)
 - uložení vytvořeného pohledu na data a jeho exportu do dalších formátů
- Pokročilý analytický software
 - příprava pokročilých ad-hoc analýz
 - export reportů do různých formátů (xls, pdf, obrázek, text, xml)
- BI a GIS
 - Statické mapové výstupy v reportech
 - Využití dat DWH v klientu GISu
- Analytický interní portál
 - integrace všech výstupů na jednom místě
 - podpora koncových uživatelů
 - personalizovaný prostor pro konkrétní skupiny uživatelů s různými potřebami

1.1.1.4.2.3 Externí přístup

- Reporting
 - přístup k reportům přes webové rozhraní
 - export reportů do různých formátů (xls, pdf, obrázek, text, xml)
 - automatická distribuce reportů
 - pokročilé řízení přístupu uživatelů k reportům i vlastnímu obsahu reportů
 - centrální správa řešení
- Ad-hoc výstupy
 - webový nástroj umožňující uživatelsky definovat výstupy z multidimenzionálních OLAP struktur

1.1.1.5 Metadata

Kritickým faktorem konzistence datového skladu a snadné orientaci koncových uživatelů ve výstupech je kvalita záznamu metadat všech částí datového skladu i výstupů, jejich následná správa a prezentace. V systému správy metadat je třeba použít adekvátní nástroj typu CASE nebo podobný nástroj a zároveň zabezpečit systém metadat metodicky – jednoznačně definovat zodpovědné osoby, které budou metadata plnit a také udržovat jejich aktuální stav. Maximum

možných údajů musí být čerpáno automaticky z jednotlivých vrstev řešení, některé údaje ale musí přesto doplnit pracovníci KÚ.

Z pohledu využití lze oblast metadat rozdělit na dvě oblasti:

Administrátor

Administrátor využívá nástrojů pro správu metadat při změnách a rozšiřování datového skladu – provádí analýzy dopadu těchto změn a zabezpečuje tak ochranu již vybudovaného řešení.

Uživatel

Uživatel má k dispozici ke každému výstupu metadata vypublikovaná na portál a jednoduše tak může zjistit, co si vlastně prohlíží – co je zdrojem dat, kdo je zodpovědnou osobou, kdy byl výstup naposledy aktualizován, jaká je metodika výpočtu, apod.

1.1.1.6 Nestrukturovaná data.

Řešení analýzy nestrukturovaných dat není na úrovni Krajského úřadu Karlovarského kraje prioritní oblastí. Z tohoto důvodu nebude v rámci projektu Datové sklady a nástroje BI tato část řešena.

1.1.1.7 Návrh datového obsahu – základní datový sklad

1. DATOVÉ TRŽIŠTĚ METADATA

Název datové struktury:	Metadata datového skladu
Popis:	Bude obsahovat základní údaje pro data, datová tržiště a kostky datového skladu. Čas aktualizace, velikost, role, dimenze, fakta, vypočítané položky, návštěvnost jednotlivých analytických aplikací, kostek, reportů, časy spuštění jobů, chyby ve zpracování dat, atd... Na kostky bude vázat názvem datové kostky, požadováno je zachování možnosti přímého odskoku z analytických aplikací na popis používaného datového zdroje, návštěvnost je nutné sledovat do úrovně konkrétního pracovníka.
Potenciální datové zdroje:	analysis services, integration services, excel ruční, aplikace pro správu metadat

Tabulka 2: Metadata datového skladu

Název datové struktury:	Metadata úřadu
Popis:	Bude obsahovat popis datových zdrojů úřadu ve zvoleném metadatovém standardu.
Potenciální datové zdroje:	aplikace pro evidenci metadat

Tabulka 3: Metadata úřadu

Předpokládaný finální počet multidimenzionálních objektů: 2

2. DATOVÉ TRŽIŠTĚ STATISTIKA

Statistické datové tržiště je průřezovým datovým tržištěm, které poskytuje referenční statistická data od externích poskytovatelů, především institucí zabývajících se sběrem a vytvářením regionálních dat. Na krajském úřadu lze najít široké spektrum činností, které jsou a budou na tržiště navázány. Určitým způsobem se tržiště překrývá s ostatními tržišti (služby).

Název datové struktury:	Souhrnné databáze ČSÚ
Popis:	KROK,MOS,Demografie,RES
Potenciální datové zdroje:	Data předávána v xls, dbf z ČSÚ

Tabulka 4: Souhrnné databáze ČSÚ

Název datové struktury:	Specifické publikace ČSÚ
Popis:	Vzhledem k neaktuálnosti informací v souhrnných databázích (KROK, MOS) je nutné využívat několika publikací ČSÚ. Nevýhodou tohoto způsobu je nadměrná pracnost přípravy importu dat, resp. i pracnost zpracování pomocí maker. Pro funkčnost a využití datového skladu je však nezbytné tato data mít a dokonce výrazně rozšířit záběr zpracovávaných publikací.
Potenciální datové zdroje:	publikace ČSÚ

Tabulka 5: Specifické publikace ČSÚ

Název datové struktury:	Regionální databáze
Popis:	V rámci statistického tržiště vznikne tzv. „regionální databáze“. Ta bude obsahovat definovaný průřez („standardizovaný set“) regionálních dat ze všech oblastí. Tato databáze bude k využití partnerům (vysoké školy) a bude rovněž tvořit základ pro regionální analýzu. Aktualizace databáze bude probíhat automaticky během aktualizace jednotlivých věcných kostek. Důraz je kladen na datový model, přenositelnost databáze.
Potenciální datové zdroje:	průřezově celý datový sklad

Tabulka 6: Regionální databáze

Název datové struktury:	Data EUROSTATU, OECD a další mezinárodní data
Popis:	V širším pojetí se jako velmi potřební jeví mít k dispozici i mezinárodní data na úrovni regionů.
Potenciální datové zdroje:	EUROSTAT, OECD, ČSÚ

Tabulka 7: Data Eurostatu, OECD a další mezinárodní data

Název datové struktury:	Data ČNB
Popis:	Měnové kurzy a přímé zahraniční investice.
Potenciální datové zdroje:	ČNB

Tabulka 8: Data ČNB

Název datové struktury:	Ministerstvo práce a sociálních věcí
Popis:	Data z MPSV. Nezaměstnanost, dávky státní sociální podpory, Průměrná mzda.
Potenciální datové zdroje:	MPSV

Tabulka 9: MPSV

Název datové struktury:	Průzkum informační gramotnosti
Popis:	Společnost STEM/MARK dodává OI šetření „Výzkum informačních zdrojů“.
Potenciální datové zdroje:	CD průzkumu

Tabulka 10: Průzkum informační gramotnosti

Předpokládaný finální počet multidimenzionálních objektů: **cca 20-25**

3. DATOVÉ TRŽIŠTĚ EKONOMIKA A ROZPOČET ÚŘADU

Některé oblasti jsou již realizovány v současném tržišti Nákladové účetnictví. Toto tržišťe by mělo být rozšířeno o další oblasti.

Název datové struktury:	Obraty – již existuje.
Popis:	rozpočtové obraty
Potenciální datové zdroje:	Ekonomický systém kraje

Tabulka 11: Obraty

Název datové struktury:	Doklady – již existuje
Popis:	obraty do úrovně rozpočtového a účetního dokladu
Potenciální datové zdroje:	Ekonomický systém kraje

Tabulka 12: Doklady

Název datové struktury:	Kapitoly, Příjmy
Popis:	Datové kostky pro potřeby analytického webu.
Potenciální datové zdroje:	Ekonomický systém kraje

Tabulka 13: Kapitoly, Příjmy

Název datové struktury:	Cash Flow
Popis:	pomocná kostka pro evidenci a výpočet cash flow.
Potenciální datové zdroje:	podklady vytvářeny na základě výstupů z kostek Obraty a doklady

Tabulka 14: Cash Flow

Název datové struktury:	Účetnictví
Popis:	Hlavní kniha, rozvaha, výsledovka, příloha, závazky, pohledávky
Potenciální datové zdroje:	Ekonomický systém kraje

Tabulka 15: Účetnictví

Název datové struktury:	Majetek
Popis:	evidence majetku

Potenciální datové zdroje:	Ekonomický systém kraje
----------------------------	-------------------------

Tabulka 16: Majetek

Předpokládaný finální počet multidimenzionálních objektů: **cca 10-15**

4. DATOVÉ TRŽIŠTĚ, VYBRANÁ SLUŽBA - DOPRAVA

Název datové struktury:	Doprava - silniční databanka, investiční akce, údržba
Popis:	Data krajské správy a údržby silnic, vazba na GIS, vazba např. na datové tržiště Integrovaný záchranný systém (dopravní nehody na silnicích), ekonomika a rozpočet úřadu, statistika, atd.
Potenciální datové zdroje:	krajský úřad, externí subjekty - ŘSD, KSÚS

Tabulka 17: Doprava

Předpokládaný finální počet multidimenzionálních objektů: **cca 5**

1.1.1.8 Návrh datového obsahu – další možná datová tržiště

1. DATOVÉ TRŽIŠTĚ „REGISTRY DATOVÉHO SKLADU“

Datové tržiště bude obsahovat 4 základní registry, které budou sloužit primárně jako (historizované) dimenze pro další datová tržiště a kostky, sekundárně jako informace o předmětné oblasti. V aktualizaci budou mít nejvyšší prioritu a budou maximálně možné podrobné. Vybrané registry bude možno v budoucnu navázat na základní registry či obdobná datová rozhraní.

Název datové struktury:	Registr územní struktury
Popis:	Bude zachycovat územní strukturu České republiky, optimálně i Evropské unie (do NUTS3 či NUTS2). Musí být zakomponovány i územní identifikace (kódy) dalších institucí, poskytující regionální data (např. ČNB, Policie ČR) včetně převodníků
Potenciální datové zdroje:	UIR-ADR, ČSÚ, základní registry, rozhraní institucí (v budoucnu)

Tabulka 18: Registr územní struktury

Název datové struktury:	Registr subjektů
Popis:	Bude zachycovat ekonomické subjekty v ČR s IČ, optimálně i bez IČ.
Potenciální datové zdroje:	ARES MF, RES ČSÚ, základní registry (v budoucnu)

Tabulka 19: Registr subjektů

Název datové struktury:	Registr příspěvkových organizací kraje
Popis:	Bude obsahovat strukturu krajských příspěvkových organizací.
Potenciální datové zdroje:	e-pusa, excel, ARES,

Tabulka 20: Registr příspěvkových organizací kraje

Název datové struktury:	Registr organizační struktury úřadu
Popis:	Bude obsahovat (historizovanou) strukturu úřadu do úrovně jednotlivých jmen.
Potenciální datové zdroje:	interní databáze úřadu

Tabulka 21: Registr organizační struktury úřadu

Předpokládaný finální počet multidimenzionálních objektů: 4

1. DATOVÉ TRŽIŠTĚ - VÝKONOVÉ UKAZATELE HODNOCENÍ EFEKTIVITY

Název datové struktury:	Výkonové ukazatele hodnocení efektivity
Popis:	Indikátory v oblasti hodnocení efektivity veřejných a interních služeb. Bude pokrývat např. Sledování metrik a alokačních klíčů procesů nákladového účetnictví, Indikátory projektů Oborové, odvětvové ukazatele výkonů
Potenciální datové zdroje:	Ruční sběr dat

Tabulka 22: Výkonové ukazatele hodnocení efektivity

2. DATOVÉ TRŽIŠTĚ VĚCNÉHO PLÁNOVÁNÍ

Název datové struktury:	Věcné plánování
Popis:	Informace o projektech a definovaných cílech a záměrů ve smyslu přípravy rozpočtu, kalkulací, čerpání, nákladovosti a řízení lidských zdrojů
Potenciální datové zdroje:	Ekonomický systém kraje, ruční vstupy

Tabulka 23: Věcné plánování

3. DATOVÉ TRŽIŠTĚ SPISOVÁ SLUŽBA

Název datové struktury:	Spisová služba (Spis, Dokument, virtuální SSL)
Popis:	Údaje o spisech a dokumentech ve spisové službě Využití zejména v oblasti správního řízení, měření lhůt, atd.
Potenciální datové zdroje:	Aplikace Spisové služby

Tabulka 24: Spisová služba

Předpokládaný finální počet multidimenzionálních objektů: 3

4. DATOVÉ TRŽIŠTĚ EKONOMIKA OBCÍ

Název datové struktury:	Ekonomika obcí
Popis:	Datová kostka z rozpočtového výkazu obcí Fin 2-12
Potenciální datové zdroje:	sběr dat, MF

Tabulka 25: Ekonomika obcí

Předpokládaný finální počet multidimenzionálních objektů: 1-2

5. DATOVÉ TRŽIŠTĚ KRAJSKÉ DOTACE

Název datové struktury:	Krajské dotace
-------------------------	-----------------------

Popis:	Datové kostky dle jednotlivých dotačních titulů.
Potenciální datové zdroje:	Aplikace KÚ

Tabulka 26: Krajské dotace

Předpokládaný počet multidimenzionálních objektů: **cca 5**

6. DATOVÉ TRŽIŠTĚ ŠKOLSTVÍ

Název datové struktury:	Školství
Popis:	Cílem je obdržet kompletní statistický popis školství v kraji + celostátní relační hodnoty. Datový obsah by měl pokrýt výroční zprávu kraje v oblasti školství.
Potenciální datové zdroje:	ÚIV, interní dotazníky OŠMS, interní databáze kraje

Tabulka 27: Školství

Předpokládaný finální počet multidimenzionálních objektů: **15 - 20**

7. DATOVÉ TRŽIŠTĚ PROVOZ

Název datové struktury:	Zaměstnanci a vzdělávání
Popis:	Data o zaměstnancích a vzdělávání. Bylo by dobré data navázat na registr organizační struktury a vytvořit určitou ucelenou podskupinu personalistiky, případně uvažovat o samostatném datovém tržišti.
Potenciální datové zdroje:	krajský úřad

Tabulka 28: Zaměstnanci a vzdělávání

Název datové struktury:	Data provozních aplikací a databází
Popis:	Jedná se požadované výstupy z provozních evidencí a software – Služební cesty, auto, tisky, atd..
Potenciální datové zdroje:	provozní aplikace a evidence

Tabulka 29: Data provozních aplikací

Předpokládaný finální počet multidimenzionálních objektů: **cca 10**

8. JEDNORÁZOVÉ DATOVÉ TRŽIŠTĚ

Název datové struktury:	Jednorázové datové tržiště
Popis:	Připravený prostor Datového skladu sloužící k jedinečným a jednorázovým ad-hoc analýzám – v DWH musí být připraveno prostředí pro možnost rychlého vytváření dočasných nových oblastí v souladu s metodikou provozu a rozvoje DWH
Potenciální datové zdroje:	krajský úřad, externí

Tabulka 30: Jednorázové datové tržiště

Předpokládaný finální počet multidimenzionálních objektů: **cca 1-5**

1.1.2 Variantní návrhy technického řešení – HW/SW/data

V oblasti **HW** se předpokládá jeho pořízení v rámci krajských technologických center.

V oblasti SW se předpokládá pořízení infrastrukturních komponent (operační systém, portál, apod.) v rámci krajských technologických center, speciální SW (databáze, SW pro podporu řešení metadat atp.) v rámci tohoto projektu.

Potenciální **Datový obsah Datového skladu** je detailně popsán v samostatné kapitole 1.1.1.7 a 1.1.1.8.

Protože řešení datového skladu neobsahuje dodávky hardwarových ani softwarových technologií a jedná se o vytvoření datového skladu, zpracování datových oblastí a zavedení portálu **nebylo uvažováno s variantním řešením**.

1.1.3 Naplnění požadavků typizovaného projektu

V souladu s typizovaným projektem je datový sklad KÚ Karlovarského kraje navržen tak, aby zajistil zpřístupnění relevantních dat na úrovni subjektů veřejné správy kraje, integraci dat z různých zdrojů, zvýšil využitelnost, výtěžnost a zkvalitnil rozhodovací procesy krajské samosprávy. Navržené řešení podporuje požadované metody uspořádání velkých objemů dat tak, aby bylo přístupné a srozumitelné uživatelům zabývajících se jejich následnou analýzou.

Řešení se skládá z dodávky softwarových komponent, vytvoření datového skladu i následných analytických a reportovacích vrstev, zpracování datových oblastí a zavedení jednotného interního a externího analytického portálu pro prezentaci dat.

1.2 Porovnání variant technologických řešení

V následujících podkapitolách jsou definována kritéria pro výběr nabídek dodavatelů a analyzována některá technická a bezpečnostní rizika.

1.2.1 Srovnání nabídek jednotlivých dodavatelů (poptávkové řízení)

Při porovnání nabídek jednotlivých potenciálních dodavatelů je třeba se zaměřit na následující skutečnosti:

- poměr mezi rozsahem zpracovaného datového obsahu a cenou,
- poměr mezi rozsahem funkčnosti řešení a cenou,
- poměr mezi cenou licencí a cenou práce ve vztahu k výše uvedeným bodům,
- rozšiřitelnost řešení s ohledem na budoucí rozvoj, robustnost a škálovatelnost komponent,
- komplexnost klientských nástrojů, přitom jejich přívětivé uživatelské rozhraní a celková ergonomie ovládání,
- reference na shodné projekty v oblasti samosprávy.

1.2.2 Výhody a nevýhody jednotlivých řešení

Při porovnávání nabídek jednotlivých potenciálních dodavatelů je nutné zohlednit nabízenou technologii i z hlediska současného know-how a zkušeností pracovníků krajského úřadu pro zajištění fungování a rozvoje řešení v provozní fázi projektu a zaměřit se tak na návratnost již vynaložených investic a celkových nákladů na vlastnictví (TCO).

1.2.3 Analýza technických a bezpečnostních rizik

Jednoznačně největším technickým rizikovým faktorem projektu je kvalita datových zdrojů. Jedná se o potřebu zachování kvality a neměnnosti formátu zdrojových dat po dobu udržitelnosti projektu především u externích poskytovatelů. Nesplnění této podmínky vede k problematickému automatickému strojovému opakovatelnému načítání dat vedoucímu od výrazného nárůstu lidské spoluúčasti při každém zpracování dat až po nemožnost načtení či propojení s historickými daty již uloženými v datovém skladu. Vhodným ošetřením tohoto rizika je smluvní zakotvení formátu rozhraní dat, jejich obsahu a periodicitu aktualizace mezi krajem a poskytovatelem dat.

Pro přispěvatele je třeba připravit takový management sběru dat, který zajistí v požadované periodicitě garantovanou průkaznost a validitu obsahu sbíraných dat. Riziko lze taktéž snížit využitím jednotné platformy pro sběr dat.

Z pohledu kvality dat může být u některých datových zdrojů problém v neaktuálnosti datových zdrojů a z toho vyplývající riziko přijetí manažerských opatření, které se mohou projevit jako chybná. Systém musí umožnit adekvátní náhrady chybějících údajů ať už z jiného datového zdroje, kvalifikovaným odhadem nebo časovou řadou atp. Z pohledu rutinního provozu je zde velké riziko, že interní uživatelé kraje nepřijmou poskytnutá data z datových tržišť jako validní, jelikož se jim může zdát, že data neposvětili, že to je záležitost IT odboru atd. S touto nedůvěrou jsou spojeny další problémy, typu jak zajistit adekvátní metodiku, odborné komentáře pro data – tedy jak decentralizovaně řídit obsah datové základny. Řešením je pomocí software nástrojů pro řízení obsahu datové základny přesun metodiky, garance za data atd. toho kterého tématu a oblasti na příslušná odborná pracoviště. Tato pracoviště se daným tématem, oborem atd. primárně zabývají a její přímo pověřeni pracovníci pak můžou provádět nejen validaci poskytovaných dat, ale taktéž ošetřovat obsah, metodiku a připojovat příslušné vysvětlující komentáře (metadata), jednoduše konstruovat potřebné analýzy atd. Je však třeba brát v potaz, aby softwarové nástroje byly pro pracovníky maximálně jednoduché, nevyžadovaly zvláštní znalosti typu SQL (MDX) dotazování atd. Touto distribucí odpovědnosti za datovou základnu na příslušné decentralizované pracoviště získají IT pracovníci prostor na koncepční rozvoj, centrální údržbu a tvorbu parametrických pohledů a reportů.

Velké riziko z technologického pohledu představuje nevhodně zvolená platforma s ohledem na další rozšiřitelnost a rozvoj řešení. Při jejím výběru je vhodné využít vlastních zkušeností krajského úřadu s dodavateli v příbuzných oblastech (databáze, portály). Zároveň je žádoucí komunikovat s dalšími krajskými úřady, které již mají s nasazováním Business Intelligence a datovými sklady zkušenosti. V neposlední řadě je pak třeba při výběru platformy vzít v úvahu míru konformity technologické platformy nově budovaného datového skladu s celkovou ICT

platformou krajského úřadu, která se projeví v celkových nákladech na její provoz a údržbu v období udržitelnosti projektu.

Z pohledu dalšího rozvoje je také třeba, aby navrhovaný systém připravil platformu pro plnohodnotný aktivní controlling ve veřejné správě (plánování opatření, sledování aktuálního vývoje a porovnávání proti plánu, vyhodnocování a přijetí nápravných opatření.)

Z pohledu bezpečnosti je nutné zajistit řízený přístup k datům a výstupům z datového skladu pro různé skupiny uživatelů. Vybraná platforma musí podporovat pokročilé a centralizované přidělování přístupových práv. Zároveň je nutné, aby část portálu určená pro veřejnost byla důsledně fyzicky oddělena od neveřejné části. Zároveň je nutné při provozování řešení dbát na precizní definování a kontrolu bezpečnostních pravidel pro přístup k výstupům a datům, což se musí zohlednit v metodikách a postupech provozu.

1.3 Doporučení a upřesnění pro účely zadávací dokumentace a realizační projektové dokumentace

Pro výběr vhodného dodavatele ve smyslu dodávky platformy a služeb je vhodné posuzovat jednotlivé nabídky multikriteriálně. Preference pouze nejnižší ceny u tohoto projektu vzhledem ke komplexnosti a kombinaci práce a SW platformy není nejvhodnější. Kromě implementace je při posuzování důležité zhodnotit i rozsahu poskytované technické podpory k nabízené ceně.

1.3.1 Specifikace zadání technického řešení

Zadávací dokumentace veřejné zakázky musí obsahovat požadavky specifikace technického řešení popsané v kapitole 1.1 a musí splňovat i podmínky SF EU – výzvy IOP č. 08.

1.3.2 Požadavky na implementaci, školení a technickou podporu

V průběhu implementace musí dodavatel zajistit:

- detailní analýzu informačních potřeb budoucích uživatelů a datových zdrojů
- návrh celého řešení s ohledem na závěry analýzy,
- instalaci a konfiguraci HW a SW prostředí,
- implementaci řešení dle návrhu,
- vypracování administrátorské a uživatelské dokumentace
- pilotní provoz a testování,
- zahájení ostrého provozu.

V průběhu školení musí dojít k:

- školení administrátorů pro zajištění správy a provozu všech instalovaných produktů,
- školení analytiků s ohledem na:

- plné využití implementovaných nástrojů,
- další rozvoj datových tržišť, tvorba vlastních pohledů apod.,
- základní uživatelské školení pro zaměstnance úřadu.
- off-line školení pro veřejnost (multidimenzionální ukázka práce)

Technická podpora na celkové řešení ze strany dodavatele musí zahrnovat garanci včasného zásahu dodavatele při nastání problematické situace - musí obsahovat garantované odezvy technické podpory a informaci, jakým způsobem a kdy lze nahlásit problém. Konkrétní rozsah a náplň je věcí nabídky dodavatele.

1.4 Provozní zajištění projektu

V následující kapitoly definují požadavky na záruku a servis a provozní zajištění SW a datových komponent.

1.4.1 Potřebné energetické a materiálové toky

Projekt nevyžaduje speciální energetické a materiálové toky v provozní fázi projektu, neboť bude využívat existující HW infrastrukturu, resp. HW infrastrukturu vybudovanou v rámci technologického centra kraje.

1.4.2 Záruky a servis

Standardní záruky na funkčnost řešení musí být požadovány minimálně na dobu dvou let. Servis bude pokryt standardní servisní smlouvou ve smyslu technické podpory. Smlouva musí být formulována tak, aby standardní údržba a provoz včetně aktualizací a dalšího rozvoje vlastními silami kraje neomezil nárok na záruku.

1.4.3 Údržba a nákladnost oprav

Technická podpora případně záruka pokrývá veškeré standardní opravy dysfunkcí a chyb. Náklady spojené s technickou podporou by měly být součástí servisní smlouvy s dodavatelem řešení. Je ale nutno počítat s tím, že po ukončení záruky bude nutno uzavřít pozáruční servisní smlouvu, což může znamenat zvýšení nákladů na provoz.

Údaje o životnostech jednotlivých zařízení

Tyto faktory na projekt vzhledem k jeho charakteru nemají vliv. Všechny relevantní komponenty mají životnost minimálně stejnou, jako je udržitelnost projektu.

1.4.4 Údaje o provozním zajištění SW a datových komponent

Vlastní provoz datového skladu představuje nové aktualizace, další vývoj a standardní údržbu. Aktualizací se zde rozumí například nahrávání dat za nová časová období do tabulek datového skladu, aktualizace OLAP objektů apod. Dalším vývojem se rozumí vývoj nových elementů v rámci

produktů definované architektury, například tvorba nových datových pump, tvorba reportů apod. Standardní údržbou a provozem se rozumí zálohování, obnovování ze zálohy, monitoring, patchování apod. Všechny tyto aktivity mohou být po dobu životnosti projektu vlastními silami prováděny vlastními silami nebo outsourcingem, pokud to bude ekonomicky výhodné.

1.4.5 Změny v provozní náročnosti vlivem opotřebení

Tyto faktory na projekt vzhledem k jeho charakteru nemají vliv. Všechny relevantní komponenty mají životnost minimálně stejnou, jako je udržitelnost projektu.