



TECHNISERV spol. s r.o.

Zapsaná v obchodním rejstříku u Městského soudu v Praze, v oddíle C, vložka 5239

Provozovna:

Moskevská 86

Praha 10, 101 00



ARIO – Software pro automatické rozpoznávání a identifikaci objektů na internetu

Software vznikl v rámci projektu „Automatická rozpoznávání a identifikace objektů v internetu a stávajících systémech PČR se zaměřením na PSEUD“ (VF20132015030).

Klíčová slova

Stahování obsahu internetu, big data, nestrukturovaná data, identifikace zájmových entit v datech.

Popis SW

Celkové softwarové řešení se skládá ze čtyř hlavních částí, jimiž jsou:

- nástroje pro sledování a stahování obsahu internetu,
- nástroje pro správu a vyhledávání uložených dat,
- nástroje pro analýzu dat,
- nástroje pro prezentaci výsledků.

Celé řešení funguje tak, že jsou automaticky sledovány vybrané internetové adresy a v případě, že se narazí na zajímavá data, jsou tato data uložena na lokální server. K datům staženým z internetu se dále přidávají data z dalších datových zdrojů, mezi které patří například emailové komunikace a dokumenty pocházející z počítačů a účtů zájmových osob. Data z tohoto servetu jsou následně analyzována a z výsledků jsou sestavovány reporty pro koncové uživatele. Celková architektura modulárního systému ARIO je ukázána na následujícím obrázku.

Sídlo společnosti:

IČO: 44264020

Bankovní spojení: KB Praha 7

Baarova 231/36

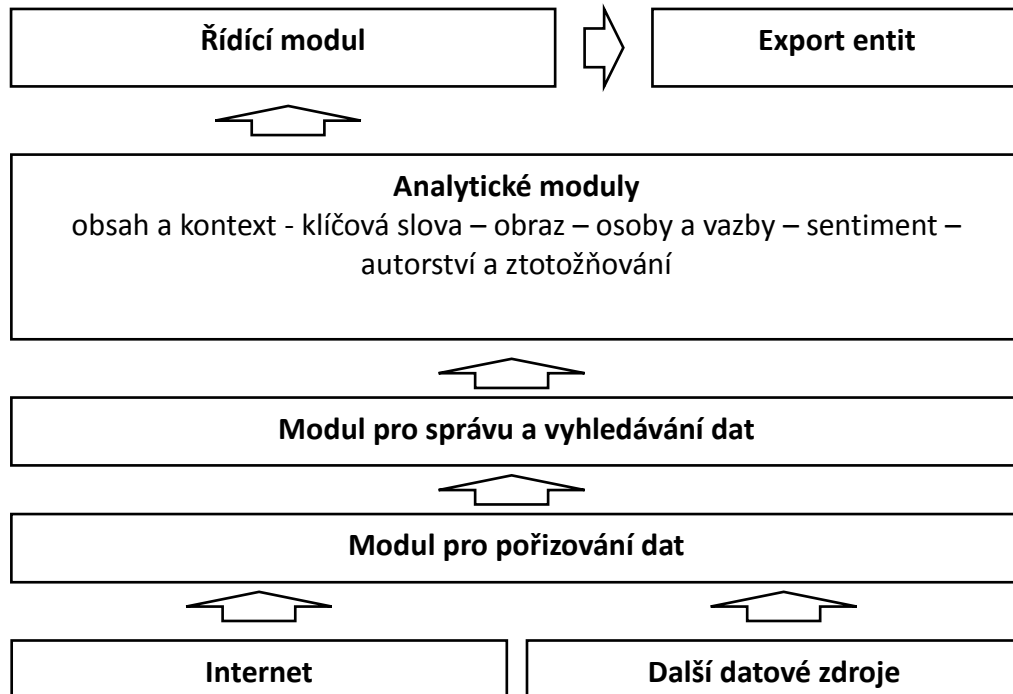
DIČ: CZ44264020

č.ú.: 435742-011

Praha 4, 140 00

www.techniserv.cz

email: info@techniserv.cz



Obr.: Architektura modulárního systému ARIO.

Obrázek ukazuje základní vrstvy a směr zpracování dat mezi moduly ARIO. Nejnižší vrstvu představuje internet a další datové zdroje, což jsou např. emaily a dokumenty z jednotlivých počítačů nebo data z pasivních sond. Tato data importuje do systému modul pro pořizování dat. Následně jsou tato data ukládána v jednotném formátu pomocí modulu pro správu a vyhledávání dat. Vedle ukládání tento modul zajišťuje i první analytickou úroveň, kterou je vyhledávání zájmových dat s cílem zúžit jejich množství tak, aby nad nimi mohly následně pracovat specializované analytické moduly. Typicky se jedná o výběr dat podle zdroje, času vzniku nebo pořízení nebo autora. Takto zúžená data jsou pak podrobována hlubším analýzám, které zajišťují specializované analytické moduly. Parametry analýz a jejich kombinace jsou řízeny řídicím modulem, který výsledky zobrazuje, ale i předává v podobě striktně definovaných entit do dalších systémů policie.

Technické parametry

Jádro software je vytvořeno v programovacím jazyce JAVA. Data jsou uložena pomocí MongoDB. Prezentační portál je postaven na technologii Liferay.



TECHNISERV spol. s r.o.

Zapsaná v obchodním rejstříku u Městského soudu v Praze, v oddíle C, vložka 5239

Provozovna:

Moskevská 86

Praha 10, 101 00



Práva k výsledkům projektu

Vlastníkem výsledku projektu je poskytovatel:

Česká republika - Ministerstvo vnitra ČR, Nad štolou 936/3, 17034 Praha 7

Zveřejnění a využití výsledků projektu se řídí ustanovením § 16 zákona 130/2002 Sb.

Kontaktní údaje

Pro bližší informace se obraťte na manažera projektu e-mail: jpatocka@techniserv.cz.

Sídlo společnosti:

IČO: 44264020

Bankovní spojení: KB Praha 7

Baarova 231/36

DIČ: CZ44264020

č.ú.: 435742-011

Praha 4, 140 00

www.techniserv.cz

email: info@techniserv.cz