Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce	Martin Mareš		
Název práce	Pseudonymizace textových datových kolekcí pro strojové učení		
Rok odevzdání	2021		
Studijní program	Infomatika	**Studijní obor**	Umělá inteligence

Autor posudku	doc. Mgr. Martin Nečaský, Ph.D. (oponent)
Pracoviště	KSI

Text posudku

Diplomová práce se zabývá návrhem a implementací uživatelské aplikace pro usnadnění procesu pseudonymizace nestrukturovaných (textových) dat, např. zápisů ze schůzek nebo soudních rozhodnutí. Řešení je navrženo a implementováno jako obecné, tj. není určené pro jednu předem vybranou doménu.

Oblast pseudonymizace je intenzivně řešena, protože se jedná o palčivý problém a častou potřebu. Autor ale v úvodu uvádí v sekci souvisejících prací pouze několik souvisejících prací. Je nutno říci, že řada dalších prací je zmiňována v dalším textu, ale systematický přehled, příp. i nějaké základní srovnání možných přístupů k aplikačnímu řešení procesu pseudonymizace v diplomové práci chybí. Nebylo to sice vysloveně cílem práce, nicméně předloženou analýzu existujících řešení pokládám za minimalistickou. Evaluaci provedenou v evaluaci rozpoznávačů pojmenovaných entit nemůžeme chápat jako evaluaci aplikační řešení pseudonymizace. Rozpoznávače jsou pouze nástrojem, byť zásadním, v rámci takové aplikace.

V části analýzy autor detailně popisuje požadavky na aplikaci včetně legislativních požadavků. Trochu matoucí je zařazení kapitoly „Návrh aplikace“ do analýzy. Je lepší rozlišovat analýzu (utřiďuji a systematicky popisuji vstupy, příp. analyzuji existující řešení) a návrh (popisuji vlastní řešení). Popis navržené architektury řešení je pouze textový, bývá zvykem ji i vizualizovat (čitelnost). Nicméně, navržená architektura je jednoduchá, tudíž to příliš nevadí. Architektura se věnuje mj. bezpečnosti. Autor zde konstatuje, že klient/server je vhodný, protože nemusíme větší části dat posílat na klienta, ale mít vše na serveru. Vzápětí ale autor popisuje, že i při plné virtualizaci existuje bezpečnostní riziko zneužití dat. To ale naopak architektura klient/server z určitého hlediska usnadňuje, neboť data jsou na jednom místě. Útočník, který využije případnou bezpečnostní „díru“ (o nich autor práce píše) tak má data všechna k dispozici. Nedalo by se uvažovat i o nějakém (třeba i fyzickém) rozdělení dat v případě opravdu citlivých dat? Rozdělení se např. provádí ve strukturovaných databázích, kdy si citlivý záznam o osobě rozdělíme uměle na více záznamů, které už tak citlivé nejsou, propojíme je proprietárními identifikátory a omezíme k nim přístup, příp. je i fyzicky rozdělíme. Toto může mít smysl nejenom pro strukturovaná data, ale i pro nestrukturovaná data.

V části evaluace rozpoznávačů pojmenovaných entit se autor detailně věnuje třem existujícím nástrojům pro rozpoznávání pojmenovaných entit v nestrukturovaných datech a jejich možnému využití pro vyhledávání citlivých informací. Takovými informacemi nemusí být jen jména osob, ale např. i názvy jiných typů entit, které mohou nepřímo identifikovat osobu nebo mohou odhalovat jinou citlivou informaci. Autor důkladně měří chování nástrojů na kolekci zápisů ze schůzek. Z výsledků měření pak dovozuje vlastní algoritmus pro vyhledávání citlivých údajů. Text je sice poměrně detailní, ale převážně se věnuje výsledkům měření, méně pak důvodům takových výsledků. Postup vyhledávání citlivých údajů, který autor dovozuje z výsledků měření, nazývá sám autor práce algoritmem, ale na to je dle mého názoru postup popsán poměrně vágně. Málo detailní je dle mého názoru i zdůvodnění, proč takový postup funguje. Nicméně, z výsledků měření funkčnost vyplývá, stejně tak lze i intuitivně pochopit navržený postup. Kapitola tak svůj účel plní.

V dalších částech pak již autor popisuje implementaci aplikace a její uživatelskou dokumentaci. Autor se příliš nevěnuje testování a evaluaci samotné aplikace. Jsou jen drobně popsány prováděné testy. S částí evaluace rozpoznávačů se ale svým detailem vůbec nemůže srovnávat. Výsledná aplikace není žádným způsobem evaluována a tak se nedozvíme, jak nakonec aplikace proces pseudonymizace usnadňuje z pohledu skutečných uživatelů. Z výsledků měření v dřívější kapitole víme, že rozpoznávání s nějakou mírou úspěšnosti funguje, ale nevíme, jak pomáhá nebo naopak nepomáhá samotná aplikace. To ale nebylo cílem diplomové práce.

Proto lze s jistotou konstatovat, že autor cíle práce beze zbytku naplnil.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

Datum 11.08.2021 doc. Mgr. Martin Nečaský, Ph.D