**ОПИС НА УСЛУГАТА**

**ПОВИК бр. 1/2015**

**НАБАВКА НА УСЛУГИ за**

**ВЕБ АПЛИКАЦИЈА ЗА АРХИВИРАЊЕ НА МЕДИУМСКИ СОДРЖИНИ**

1.1 **Веб апликација за архивирање на содржини од медиумите**

Главната цел на веб апликацијата за архивирање на содржини од медиуми е да овозможи правење на „слика“ (snapshot) од веб страница со временска ознака кога е земена и зачувување на истата на заштитен сервер. „Сликата“ на веб страницата треба да овозможи документирање на содржината на одредена веб страна (со мултимедијалните содржини во неа), заедно со нејзиниот целосен изглед во точно определн временски момент, и тоа во две форми

1. Запис во форма на текст (HTML) и слики, кој ќе се зачува во база на податоци;
2. Снимка на екранот, односно изгледот на страницата во графички формат (JPG и/или PNG)

Со развојот на оваа апликација ќе се овозможи дополнителна алатка во имплементацијата на методологијата на Сервисот за проверка на факти во медиумите каде се бара рецензентите да направат „слика“ од веб страницата која е рецензирана во моментот. Иако рецензиите содржат и линкови до самата веб страна, правењето на слика од веб страната е потребно бидејќи во меѓувреме содржините на самата страна може да бидат сменети. Ова често се случува со сомнителни написи кои содржат грешки или манипулации. Вклучувањето на овие „слики“ овозможува доказ на кој се поткрепува рецензијата и е наменет за читателите, како и за заштита при случај на тужба од рецензираниот медиум. На овој начин, Сервисот за проверка на факти служи како дополнителна јавна архива за манипулациите во медиумите.

Веб апликацијата за архивирање на содржини од медиумите ќе биде хостирана на дедициран сервер (dedicated server) надвор од Македонија и ќе овозможи простор за зачувување на содржините од веб страниците кои се селектирани од корисници на два начина:

* Рачно – според процедурите кои веќе се имплементираат во правењето на рецензија, корисниците на апликацијата ќе можат да прават слики на веб страницата со додавање на веб локацијата на написот. Овие слики кои се снимени со апликацијата ќе бидат јавно достапни и секој корисник може да се повика на нив со линк или да ги вметне во некоја содржина. Овие можности најпрвин ќе го намалат времето за изработка на рецензија, но исто така ќе служат и на сите корисници како практична алатка и како начин да се привлечат кон другите содржини на Сервисот вклучувајќи се во дискусија за неговите содржини.
* Автоматски – прес клипинг од важни медиумски сајтови. Типот на Сервисот за проверка на факти преку апликацијата ќе избере листа на веб сајтови од кои
* редовно и периодично ќе се собера содржината и ќе се сними на сервер. Од еден напис ќе се направат неколку слики во одреден временски интервал со цел да се забележат промените во самиот напис за тој временски период. На пример, апликацијата ќе проверува за нови содржини од одреден веб сајт на неколку минути, и доколку идентификува нов напис ќе земе слика од него во тој момент, после неколку часа, следниот ден и по неколку дена.

Развојот на овие два типа функционалности на веб апликацијата ќе овозможи таа да функционира како македонска верзија на archive.org, документирајќи написи и овозможувајќи ги како јавен запис од медиумската продукција корисни за новинарите како и за општата јавност. Секоја слика ќе биде поврзана со временска ознака – информација за датумот и времето на креирање на написот, со што било кој корисник ќе може да види како одреден напис изгледал во одреден период од времето.

1.2. **Веб апликација за споредување на написи и детекција на плагијаторство**

Плагијаторството добива нови домензии он-лајн бидејќи содржините може лесно да се копираат и користат. Новинарите понекогаш копираат делови од написи директно од соопштенија за печат или ги копираат целите написи од изворот без никакви или со мали измени. Додека е разбирливо понекогаш да се цитираат соопштенијата за печат и да се користат различни извори, користењето само на владините или корпорациските соопштенија и копирањето без наведување извор претставува сериозен проблем. Проблемите со “copy-paste” новинарството кои вклучуваат целосно копирање на содржини меѓу интернет портали преку кои овие стануваат средство за дистрибуција на манипулации, спинови и грешки, се идентификувани повеќе пати во анализите, рецензиите и новинарските лекции објавени од експертите преку Сервисот за проверка на факти во медиумите.

За среќа иако плагијаторството е многу лесно да се изврши on-line, со соодветни алатки исто така може лесно да се пронајде првобитниот извор, или барем различните членови во низата на копирани содржини. Со цел на поддршка на културата на „споделување на наведување извори“ и придонесувањето во зголемувањето на професионалните и етички стандарди во новинарството во Македонија, Сервисот за проверка на факти ќе постави алатка со која ќе може да се споредуваат написи објавени онлајн од најголемите македонски веб портали и да се споредат нивните сличности или разлики статистички и визуелно разбирлив начин. Исто така визуелно ќе биде претставена хронологијата на објава на написи почнувајќи од изворот (каде е прв пат објавен) до последната негова верзија со нагласување на промените на процентуален начин.

Развојот на оваа апликација бара користење на принципите на data mining и визуелизација, вршејќи ги следните функции на text mining

1. Автоматско пребарување и симнување написи од македонски онлајн портали за вести, преку директно поврзување со веб апликација за архивирање на содржини од медиумите од точка 1.1.
2. Автоматско идентификување релевантни информации (текст, клучни зборови, датуми) од овие преземени документи за цел на понатамошна анализа на содржините
3. Создавање онлајн архива која ќе чува податоци за промените на обработените написи во определен временски интервал
4. Анализа на трендови (на пр., тренд на зборови, настани и сл.) прикажано визуелно на едноставен и разбирлив начин.
5. Анализа на сличности меѓу написи (на пр., одредување дали написот е можен плагијат од друг напис, со процент на разликување и подредување на сличните написи по датум и час на објавување) прикажано визуелно, на едноставен и разбирлив начин.
6. Други автоматски анализи кои ќе се сметаат како додадена вредност на системот, како на пример одредување пристрасност (sentiment analysis).