Úkol č.2 do předmětu ISKM 55

Kadaňková\_Kateřina\_449127

Úkol se věnoval tématu x\_path, které se probíralo na předešlé hodině.

V prvním úkolu jsme dostali zadáno vytvořit v Google Sheets dynamický seznam, který by obsahoval TOP 48 nejlepších fitness věcí, které jsou na webu heureka.cz. Hned první bod úkolu mi dal poměrně dost zabrat, protože jsem nebyla schopna vygenerovat na jeden x\_path příkaz celý seznam všech 48 produktů. Nakonec jsem po hodině snahu vzdala a vyřešila jej následovně. Každá stránka nabízela právě 24 produktů, tak jsem seznam rozdělila a pomocí dvou různých příkazů odkazují na jiné url odkazy seznam na dvakrát nahrála. Nebyla jsem si úplně jistá, zda pouhý seznam názvů by se dal nazývat dynamickým, a tak jsem k názvům vytvořila i druhý sloupec odhazující na url odkazy jednotlivých produktů. Dalším požadavkem bylo připojit k vyjmenovaným produktům i cenové rozpětí, za které je daný produkt možné koupit. Tento příkaz byl poměrně jednoduchý a zvládla jsem jej bez větších obtíží. Třetím podúkolem pak bylo připojit ještě hodnocení jednotlivých produktů, a to se opět ukázalo být poměrně velkým oříškem, který se mi bohužel nepodařilo vyřešit. Vygenerovat hodnocení byla poměrně jednoduchá záležitost, ale problém nastal v okamžiku, kdy některé produkty neměly hodnocení uvedené a v tom případě se hodnocení jednotlivých produktů při importu přeskupila a v některých případech tak byla přidělena k nesprávným produktům. Zkoušela jsem tento problém vyřešit přiřazením k hodnocení i počty recenzí, které tak v některých případech skutečně suplovaly údaj o hodnocení. Jak se ale ukázalo, ani toto řešení nebylo 100 %, protože některé produkty měly právě 0 hodnocení, a tím pádem problém stále přetrvával. ☹

Odkaz na tabulku: https://docs.google.com/spreadsheets/d/18KgBJWVqj4DbCUjb-Zv3vICcDGPD7JkeMHpuxzLjOO4/edit?usp=sharing

Druhý úkol mi pak přišel mnohem zábavnější, protože nám představil nástroj <https://demo.geneea.com/>, který provádí analýzu nad textem. Úkolem bylo vybrat delší článek, a ten pak nechat nástrojem zanalyzovat. Příjemně mě překvapilo, kolik různých hodnot analýza vyprodukovala. Zaujal mě údaj o sentimentu v textu, poté se mi velmi líbil entity cloud, který přehledně vyobrazuje nejčetnější slova v textu a čtenář má tak možnost během pár sekund vstřebat, o čem daný článek pojednával. Kromě toho pak nabízí i seznam odkazů („tagů“) a témat, o kterých článek pojednává a pro přehlednější orientaci na ně v textu přímo odkazuje. Následně analýza přináší seznam entit, které je možné v textu identifikovat, jsou jimi např. jména osob, lokací, organizací, … Do analýzy je možné se ponořit i trošku hlouběji skrze zkoumané entity a jejich vzájemné vztahy. Ale tato část už mi nepřišla pro mě jako amatérského čtenáře až tak zajímavá. Jinak si určitě dokážu představit, že kdybych psala nějaký text a chtěla bych se trošku zorientovat v relevantní literatuře, tento nástroj by mi to určitě usnadnil. ☺

Závěrem jsme si měli vyzkoušet nástroj <http://voyant-tools.org/>. Já už jsem se s nástrojem setkala, proto pro mě dané rozhraní nebylo úplně neznámé. Nahrávala jsem do něj kdysi seminárku, abych získala entity cloud, který by přehledně shrnul klíčová slova z textu, a na to mi nástroj přišel super. Pak mě samozřejmě upozornil na často opakovaná slova, kterých bych se mohla vyvarovat, na základě toho jsem pak některé pojmy nahradila pojmy synonymními. Obecně mi tyto nástroje přijdou užitečné při psaní delších textů, když se člověk už sám nedokáže přehledně orientovat v textu a správně vnímat např. již zmíněné často opakující se slova, protože daný text je zkrátka už příliš obsáhlý.