Blogový článek

1 část: Jak jsme postupovaly?

Mám data! Ale to jsme ještě nevěděly, co nás čeká…

Vyzkoušet si datovou analýzu na konkrétním případu, je super věc, ale má to malý háček. Musíte mít data. Zkoušely jsme hledat v oblastech, které by nás zajímaly, ale sehnat data, která by nebyla příliš agregovaná, se pro naše účely nenašla. Pak jsme objevily úřad duševního vlastnictví a slibný odkaz s názvem data pro strojové zpracování. Zkoumat vynálezy, znělo zajímavě, tak jsme se do toho pustily.

Získání dat

Úřad duševního vlastnictví poskytuje datové sady ve formátu \*.xml, kdy publikace probíhá formou plného exportu databáze a následných přírůstků. Znělo to skvěle, stáhnout pár souborů a máme hotovo! ALE! Data nelze vybrat na základě časového období, které chceme zkoumat a musíme stáhnout vše. No, jenže každý zip soubor, obsahuje asi 100 složek se 100 \*.xml soubory a k tomu plno obrázkových příloh. Paměť PC nestačí – musíme na to chytře přes Python. Stáhni soubor zip, který extrahuj, ale do složky nahraj jen  xml soubory. Funguje!

Máme celkem 1 737 619 souborů ve formátu \*.xml. Přichází na řadu porada s Chat GPT, který poradí, že \*.xml soubory se parsují s pomocí knihoven xml.etree.ElementTree, na řadu přichází i pandas či knihovny csv a os. Parsujeme data s podmínkou, že datum přihlášky je větší 2013 a výsledné vybrané hodnoty ukládá do \*.csv souboru. Data následně čistíme a mažeme přihlášky, které nepodaly veřejné vysoké školy. Vzhledem k tomu, že variant zápisu škol je tolik, je nejrychlejší úprava v Excelu. Následně bylo třeba tabulku normalizovat a zbavit se některých duplicit, pro které byly vytvořeny vazební tabulky. Jeden patent totiž mohlo podat více spolupracujících škol a týkají se více oblastí zájmu. Pro dočištění dat jsme použili Azure Data Studio. Jedna faktovka byla hotová.

Trápení pokračovalo s parsováním dat z \*.pdf souborů. Zkoumáme totiž licenční příjmy univerzit, jejichž hodnotu získáme z výročních zpráv vysokých škol, které jsme ručně postahovaly z webových stránek vysokých škol. Všechny výroční zprávy obsahují řádek s hodnotou licenčních příjmů v tabulce označené jako A1. Pro extrakci dat, jsme zkoušely použít knihovny PyPDF2, PDFplumber i jiné. Na řadu přišel i REGEx, se kterým nám pomohla mentorka Adéla. Nicméně po 8 hodinovém trápení na prvním Hackatonu se všechna data nepodařilo stáhnout. Vylepšováním skriptu se něco podařilo, ale i tak jsme část byly nuceny naházet ručně.

Po prvním Hackatonu jsme měly již každá dost dat ve svých laptopech a vyvstala potřeba kooperovat a verzovat. Z nabízejících se možností jsme vybraly GitHub, odzkoušený již s domácích úkolů, který se jevil jako v klidu. Po hlubším zkoumání branchí se nám na druhém hackatonu však nepodařilo rozjet jejich system a tak jsme přispívaly do mainu. Pro zvědavé možno shlédnout zde(<https://github.com/bejka2012/Patenty_BI>).

Data jsme nahrály do Power BI a abychom si nemusely data ukládat lokálně, nechaly jsme datové sety v Azure Data Studio databázi, což se osvědčilo v momentě, kdy jsme ještě objevily nějakou nesrovnalost v datech a v Azure Data Studio pomocí jazyka SQL upravily data. Následný update v Power BI pak byl jen otázkou chvilky.

Datový model

Datový model jsme si původně vyrobily dle našich představ o tématu. Po hlubším seznámení s datasetem jsme udělaly patřičné úpravy níže uvádíme hotový model.

2 část: K čemu jsme došly?

Naše výsledky výzkumu

Nás projekt byl zaměřen na inovace na veřejných vysokých školách v posledních 10 letech. Chtěly jsme zmapovat, kdo patenty přihlašuje, které školy, jestli se dají výsledky výzkumu aplikovat nebo jsou spíše teoreticky zaměřeny.

Trocha teorie

Pokud si chcete váš vynález nechat patentovat, musíte se jednat o **technické řešení**. Jinak řečeno, za návod k technickému řešení, které má být dosaženo technickými prostředky. **Patenty** se udělují na vynálezy, které jsou nové, jsou výsledkem vynálezecké činnosti a jsou průmyslově využitelné. Maximální doba ochrany patentu je 20 let.

Pro své technické řešení je možné si zvolit ochranu **užitným vzorem**. Užitný vzor někdy bývá považován za jednodušší variantu vynálezu. Oproti patentové ochraně je ochrana užitným vzorem rychlejší a méně nákladná. Maximální doba platnosti užitného vzoru, při placení poplatků za prodloužení platnosti zápisu, je oproti platnosti patentu poloviční, tj. 10 let (4 roky a můžete si dvakrát prodloužit ochranu na celkem 10 let).

?Evropský patent?

Patenty na školách

Jelikož patent či užitný vzor musí být technickým řešením, není překvapením, že nejvíce přihlášek podávají univerzity ČVUT, VŠB-TUO a VUT. Překvapením je například Akademie múzických umění v Praze s celkovým počtem 12 přihlášek (jedním z nich je například zařízení ke sledování provozu varhan).  Celkem 4 univerzity nepodaly žádnou přihlášku Akademie výtvarných umění v Praze (AVU), Janáčkova akademie múzických umění (JAMU), Slezská univerzita v Opavě (SU), Vysoká škola ekonomická v Praze (VŠE).

Školy využívají ochranu vynálezů především pomocí užitného vzoru (62%) než patentů či evropského patentu (celkem 38%), což je pro školy levnější. Celkem bylo podáno 4 339 přihlášek, z nichž jen 233 nebylo uznáno a patent nebyl udělen, což poukazuje pouze na 5% úspěšnost žadatelů.

? Info o výši poplatků

Protože s patenty souvisí povinnost hradit udržovací poplatky, část patentů je ukončena předčasně. Za zkoumané desetileté období se jednalo o 620 patentů, což představuje České vysoké učení technické V Praze - 13 %, Vysoká škola báňská - Technická univerzita Ostrava - 15%, Vysoké učení technické v Brně 17%. Nejvyšší míru vykazuje Akademie múzických umění v Praze 33%, ale i z důvodu že má pohých 12 přihlášek.

? graf

Majitelé přinášeli inovace především v oblastech fyziky, chemie, zdraví, záchrany života, zábavy, tváření, elektřiny či staveb.

Licenční příjmy univerzit

Nepodařilo se nám získat data za roky 2014 - 2016 z Českého vysokého učení technického , i když jsme oslovily univerzitu přímo. Univerzity nemají povinnost zveřejňovat výroční zprávy o hospodaření starší 5 let na svých webových stránkách a do uzávěrky projektu nám data bohužel neposkytli.

Své prvenství si drží České vysoké učení technické v Brně také ve výši licenčních příjmů v absolutním vyjádření za sledované období. V těsném závěsu se umístila Mendelova univerzita v Brně a Vysoké učení technické v Brně. Můžeme sledovat, že od roku 2016 se univerzitam velmi daří a licenční příjmy mají rostoucí charakter. Celkem za 10 leté období všechny univerzity tj. 26, měly příjmy z licenčních poplatků ve výši 193 mln.

Průměrný roční příjem všech univerzit je 744 485 CZK. Nejvýkonnější, již zmiňovaná univerzita ČVUT, průměrně za rok inkasuje více než 3,5 mln CZK, VÚT jen o něco málo a to průměrně ročně 3,1 mln CZK dále Masarykova univerzita 2,6 mln CZK. Jelikož medián příjmů všech univerzit je o mnoho nižší než průměr a to jen 146 500 CZK dovedlo nás ke kategorizaci příjmů. Rozdělili jsme příjmy do tří kategorii přičemž hranice jsme určili přibližně podle již zmíněného mediánu a průměru. Největší variabilitu příjmů pozorujeme v kategorii Hight Income, kde hodnoty příjmů sahají od středních milionových částek až po extrémní hodnoty (outliery) přesahující 5 mln. Do této kategorie spadá Technická univerzita v Liberci, která v roce 2014 dosáhla příjmů z licencí 6,3 mln či Jihočeská univerzita v Českých Budějovicích, která v roce 2023 měla příjmy z licencovaných patentů 5,3 mln.

Poskytování licencí je jedním z prostředků k získání finančních příjmů nebo jiné formy prospěchu z výsledků vynálezecké činnosti či předmětů průmyslového vlastnictví (patent, užitný vzor atd.). K realizaci této komercializace průmyslového práva a duševního vlastnictví slouží licenční smlouva. Licencí se tedy poskytuje nabyvateli licence právo k výkonu průmyslových práv, tj. například vynález chráněný patentem vyrábět a obchodovat s vyrobenými výrobky. Toto právo, například patent, se tedy licencí neprodává, ale zůstává i nadále v majetku poskytovatele licence. Ve výroční zprávě nelze zjistit, z kterých konkrétních patentů plyne daný příjem, nicméně nějaký přehled máme z databáze patentů.

Držitel patentu se může rozhodnout zda bude svůj patent licencovat či nikoliv. Důvodem pro ponechání si patentu bez licence je získání konkurenční výhody, udržení si monopolu nad výrobou či distribucí produktu nebo bezpečností či etické důvody. Naopak pro přístup

Vysoký počet patentů nemusí nutně znamenat vyšší licenční příjmy co vyjadřuje slabou závislost mezi počtem licencovaných patentů a příjmy. Mendelova univerzita v Brně a Západočeská univerzita v Plzni dosahují významných příjmů i s relativně nízkým počtem patentů. Nejúspěšnější v komerializaci svých výzkumných činností je již dříve zmiňované České vysoké učení technické v Praze. Problemy s efektivní monetizací vidíme u Vysoké školy báňské - Technické univerzity Ostrava.

V danem období bylo licencováno celkem 104 patentů podaných v letech 2014-2023, z toho 23 CVUT, 22 VSB=TUO, prekvapive 11 UTB (z celkově 199 přihlášek) a VUT 9.

Počet licencovaných patentů v průběhu času klesá, dalo by se tedy říci, že patnetnované vynálezy nemají prozatím pro firmy nazajímavé.

Musíme ještě podotknout, že české veřejné vysoké školy získávají příjmy také smluvním výzkumem s podniky či jinými institucemi. Tyto výzkumy bývají laděné spíše na míru konkrétnímu subjektu, s kterým spolupracují a pro koho vyvíjejí. Při zběžném pohledu do výročních zpráv jsme zjistily, že smluvní výzkum je pro vysoké školy zajímavější s ohledem na daleko vyšší příjmy z nich plynoucí.

Náš závěr je, že ačkoli pro školy může patentování výsledků výzkumu přinášet jistou prestiž a uznání jako inovátoři v daných oborech, není pravděpodobně toto pro školy hnacím motorem. Více se soustředí na aktivity, které přinášejí vyšší ekonomický užitek. Pokud chcete studovat technickou školu, může být pro studenta zajímavým zjištěním, jestli se škola věnuje inovacím a jestli se za dobu studia dostane k technologiím budoucnosti.