

S PYTHONEM PO STOPÁCH ZANIKLÝCH KNIHOVEN NA 1. LÉKAŘSKÉ FAKULTĚ UNIVERZITY KARLOVY



1. LÉKAŘSKÁ
FAKULTA
Univerzita Karlova

PhDr. Mgr. Markéta Ivánková, Ph.D.
Ústav vědeckých informací

O čem bude řeč?

- Jak může programování v Pythonu pomoci při provenienčním průzkumu a proč se vyplatí dát *scriptům* šanci?
- Staré tisky z berlínských lékařských knihoven, nyní ve fondu 1. LF UK
- Digitální vhled do roztržitých, resp. neúplných dat
- Rekonstruovat souslednost přesunů stovek exemplářů knih



Proč?

množství údajů

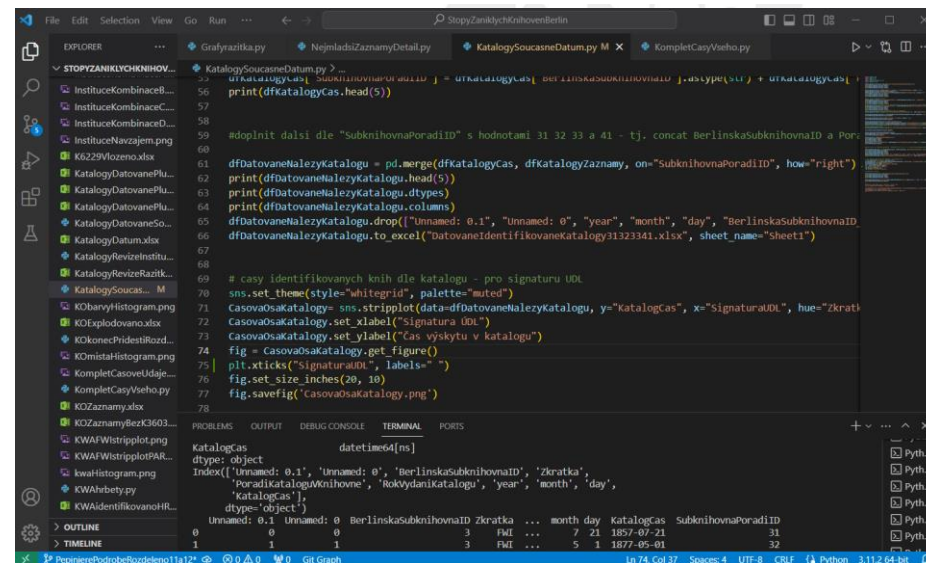
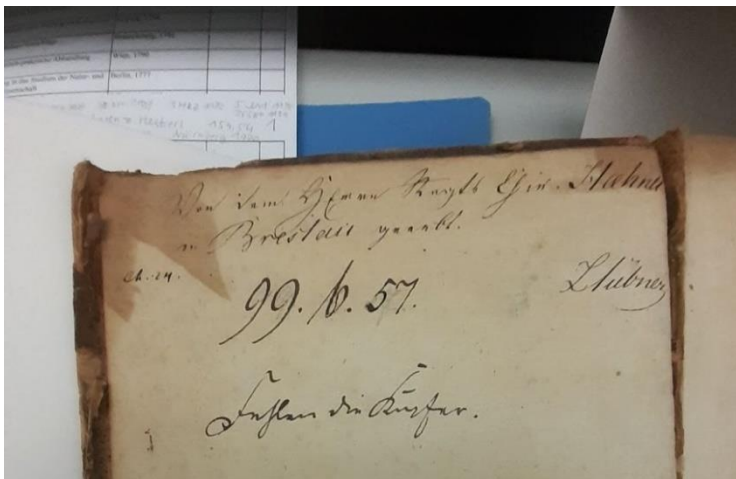
neúplných, ne vždy jistých

proto počítač

průhledná syntaxe

a řada knihoven pro vizualizace

proto tentokrát Python

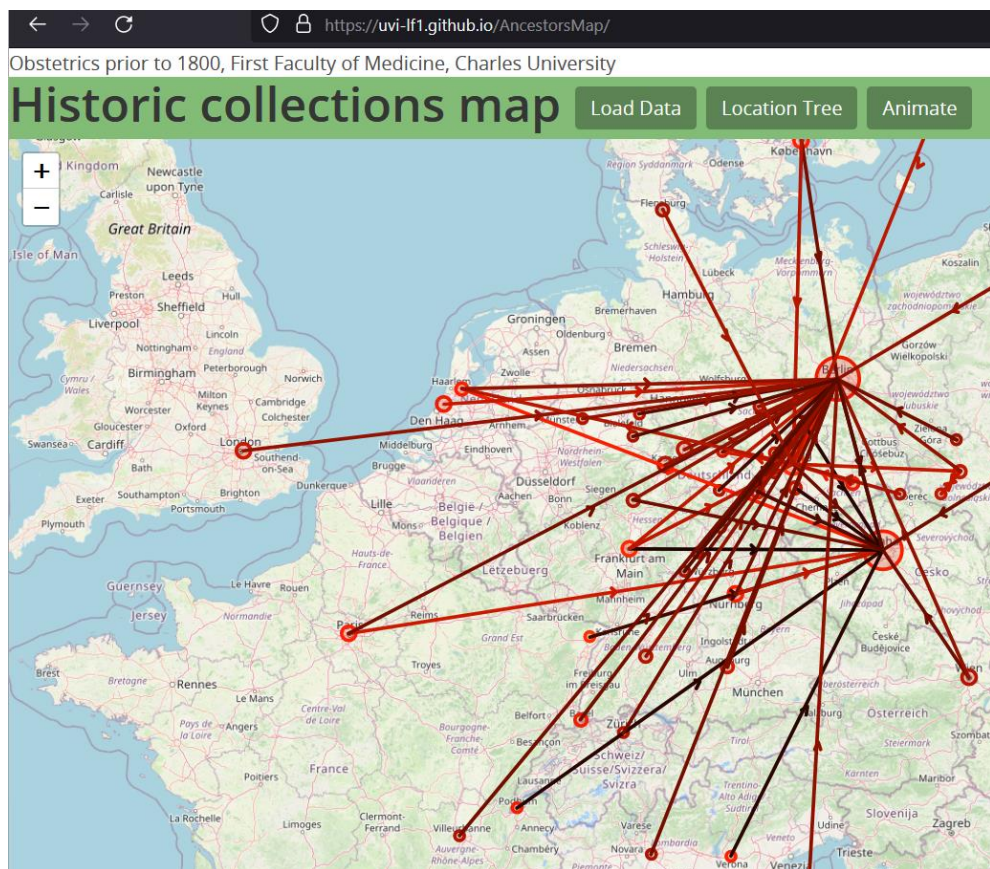


Sbírka STT ÚDL 1. LF UK:
> 2600 svazků

Data zatím z 816 svazků:
z toho 299 „z Berlína“

> (900+300) prov. záznamů
+ zvlášť zpracovaná data
o označení hřbetů





<https://uvi.lf1.cuni.cz/historicke-fondy/historic-collections-map>

Proč najisto po Berlínu?

2020

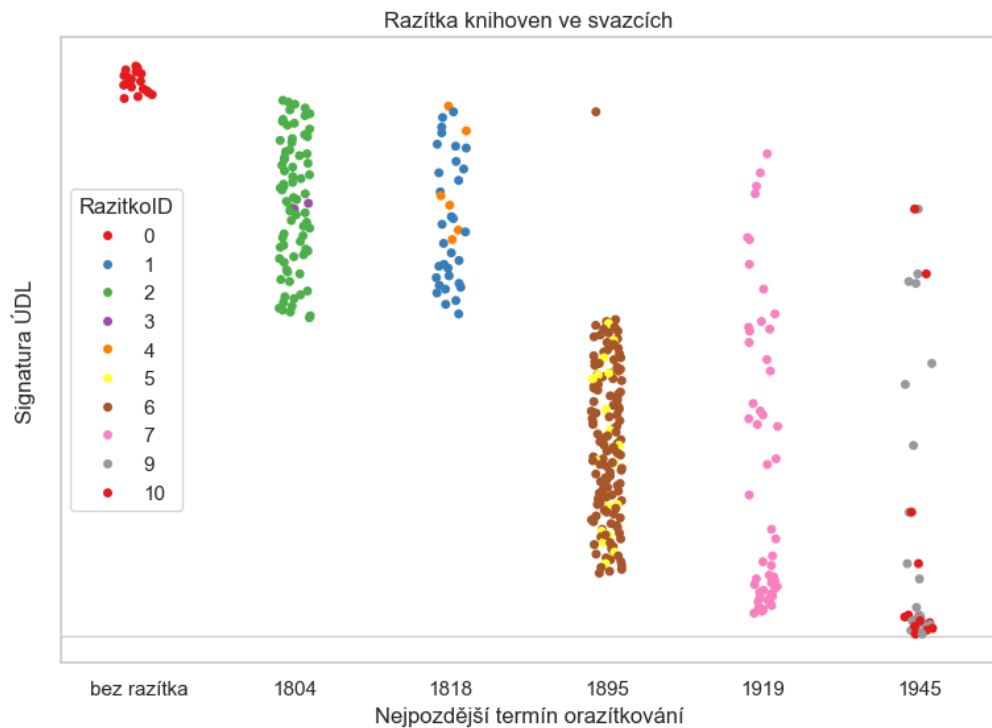
Gynekologicko-porodnická literatura ve sbírce starých tisků
Ústavu dějin lékařství 1. LF UK

Tematicky ohraničené (146 svazků)
Zjistila jsem, že 2 velké disjunktní soubory



Foto: archiv ÚVI

Poměrně přímočaře razítka institucí a čas



Zdroj: vlastní výpočet a sběr dat

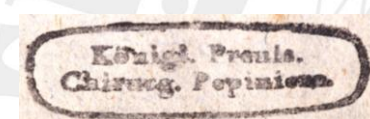
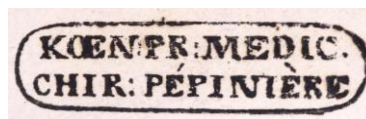
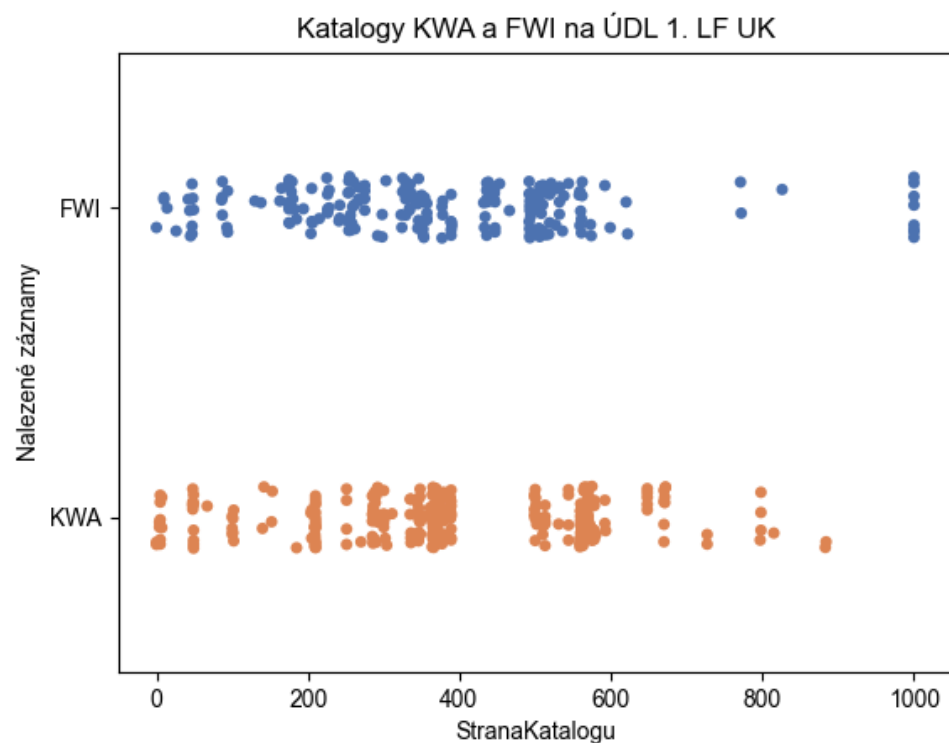
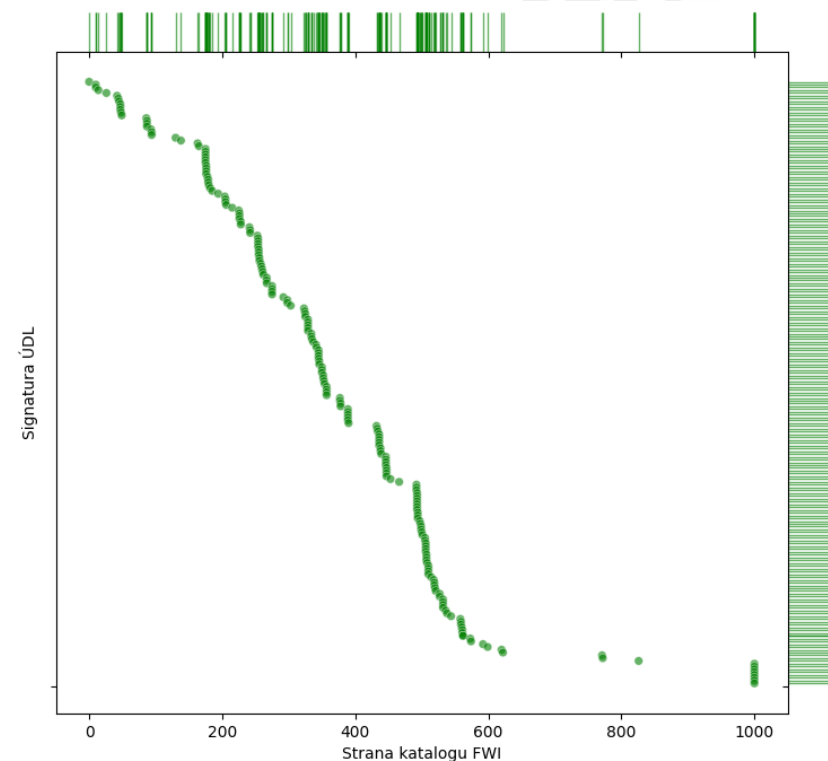


Foto: archiv ÚVI

Párování signatur a údajů o titulu v býv. katalogích



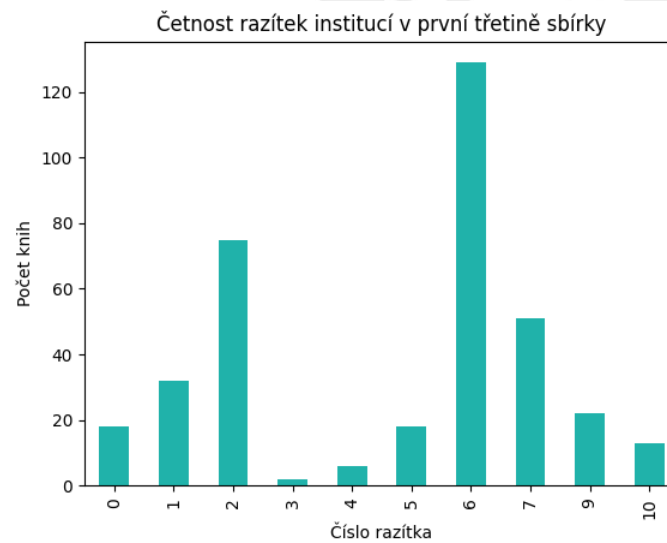
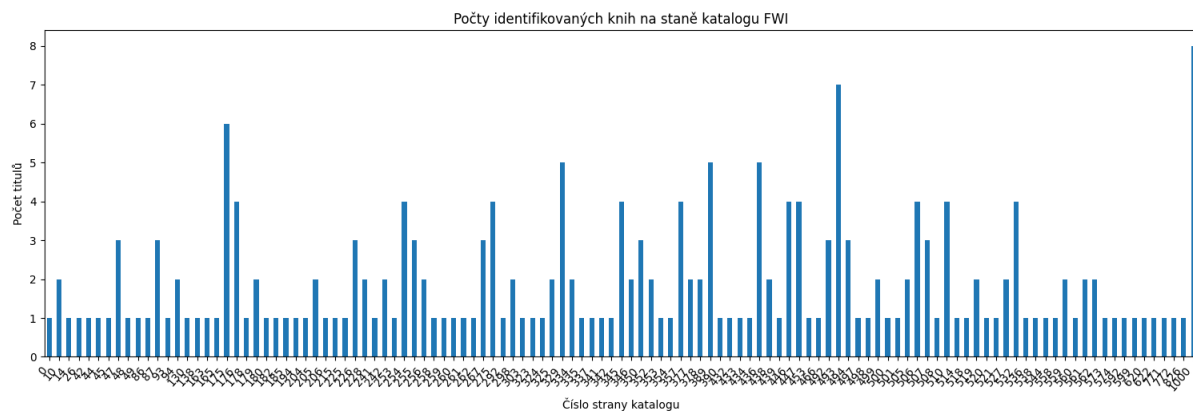
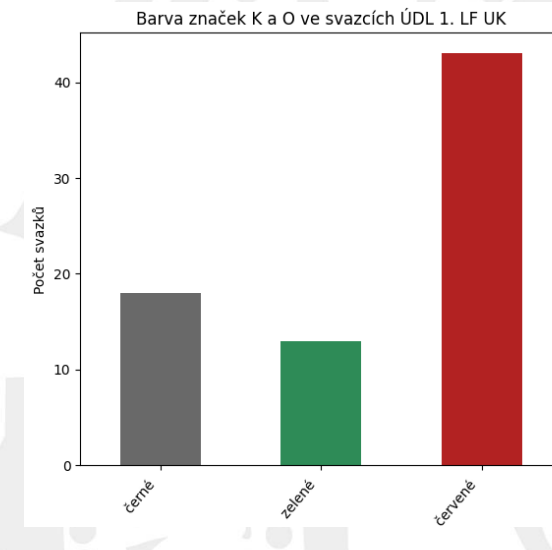
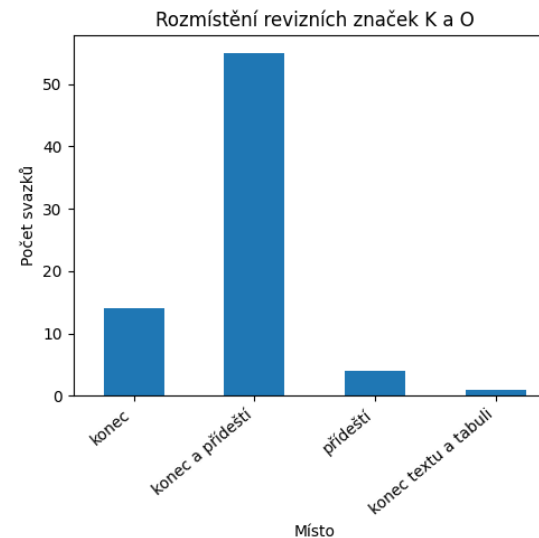
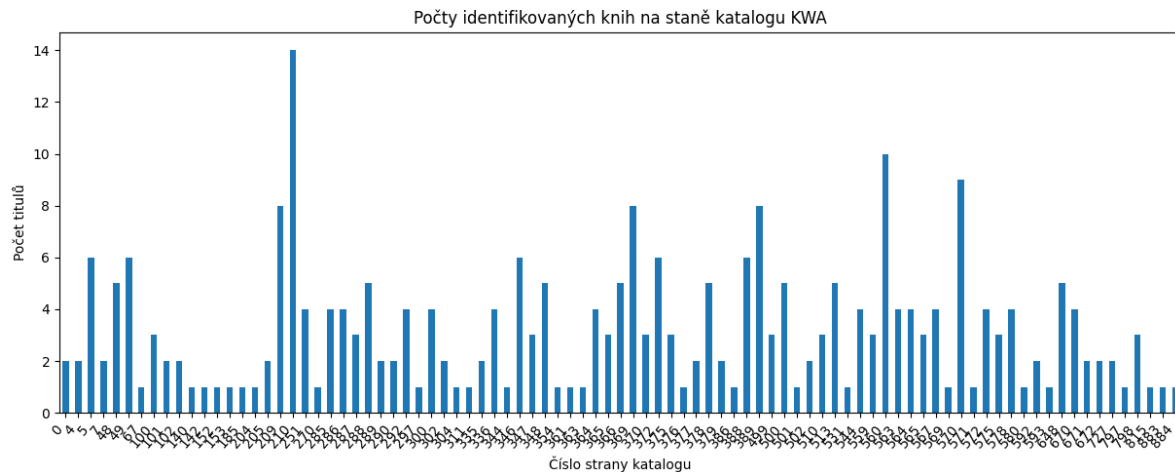
Zdroj: vlastní výpočet a sběr dat



Zdroj: vlastní výpočet a sběr dat

- *Katalog der Bibliothek des Koeniglichen medizinisch-chirurgischen Friedrich-Wilhelms-Instituts*. Berlin: Druck der Gebr. Unger'schen Hofbuchdruckerei, 1857. + dodatky 1877 a 1883.
- *Verzeichnis der Büchersammlung der Kaiser Wilhelms-Akademie für das militärärztliche Bildungswesen*. (Dritte Ausgabe). Berlin: Verlag von August Hirschwald. Unter den Linden 68, 1906.

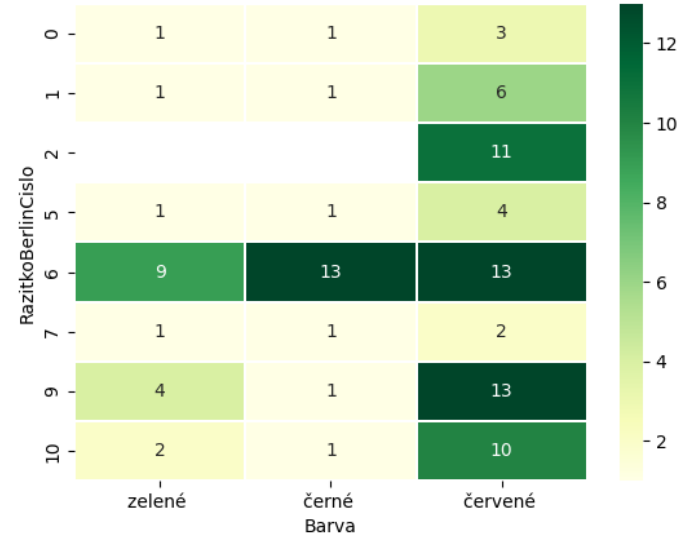
Histogramy snadno a rychle



2 sv. dodatků

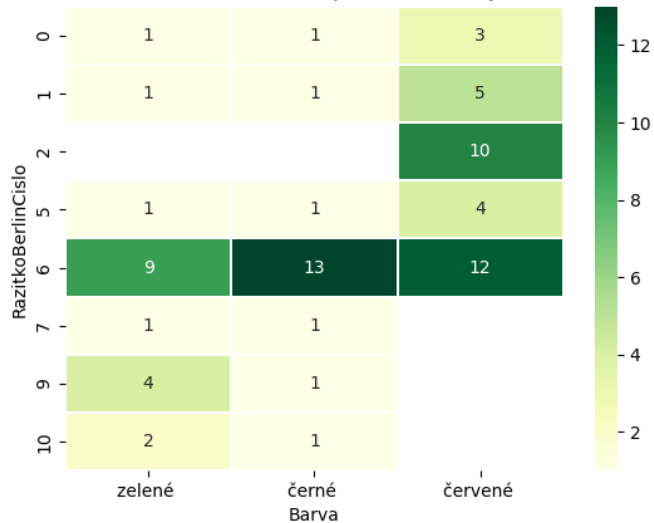
Zdroj: vlastní výpočet

Heatmap razítka instituce - barva revizního vsehny vyskyty

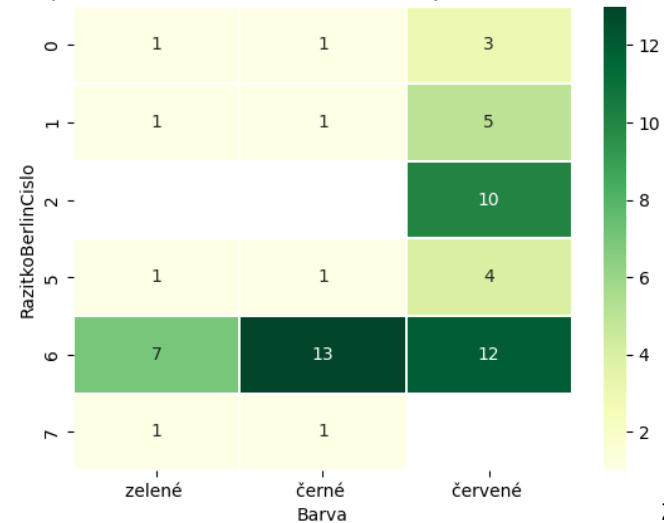


- Příklad často se objevujících červených značek současně s razítky Deutsche Ärzte-Bücherei
- Přibližně 1/2 svazků s červenou značkou na konci textu + zadním přideštlí je ve svazcích označených DÄB, zbytek se jí přisoudit nedá

razítka instituce - barva revizního, po odebrání nejmladšího červeného 9/10



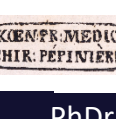
heatmap razítka instituce - barva revizního, po odebrání všech 9/10



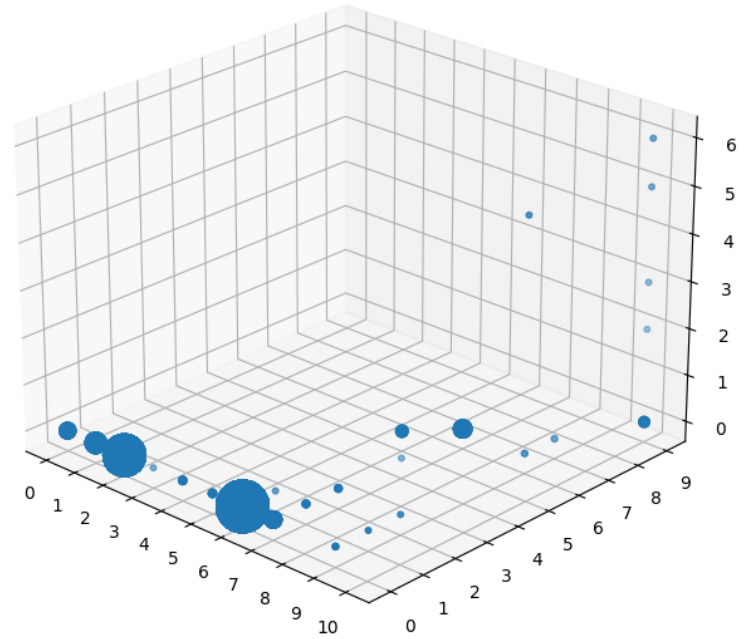
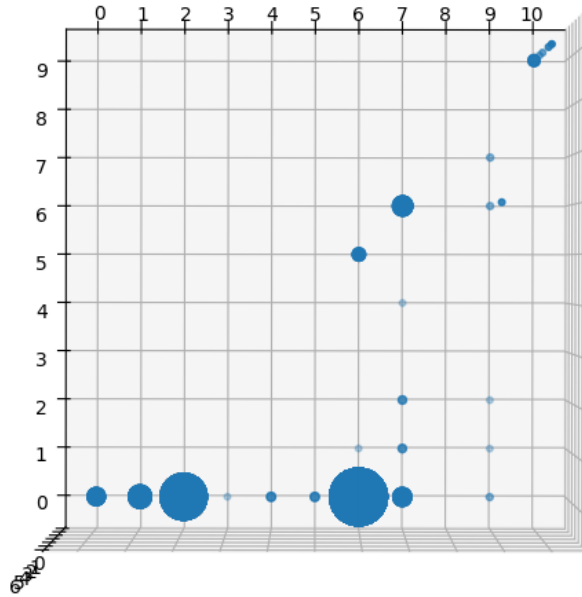
Zdroj: vlastní výpočet



Razítka institucí společně 3D

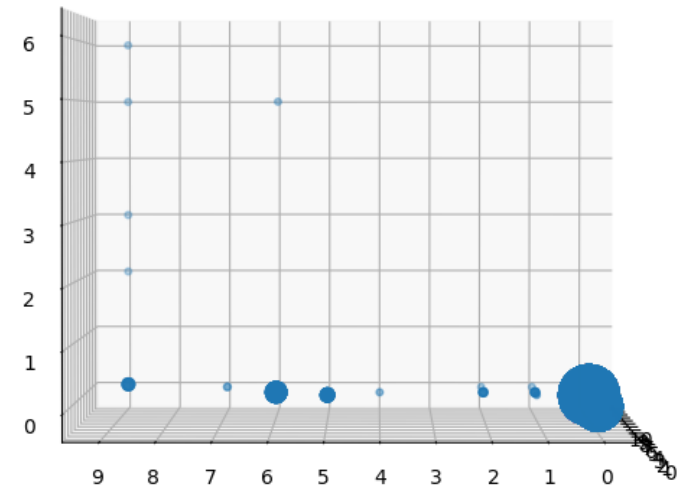


Razítka institucí společně 3D



velikost ~ počet případů

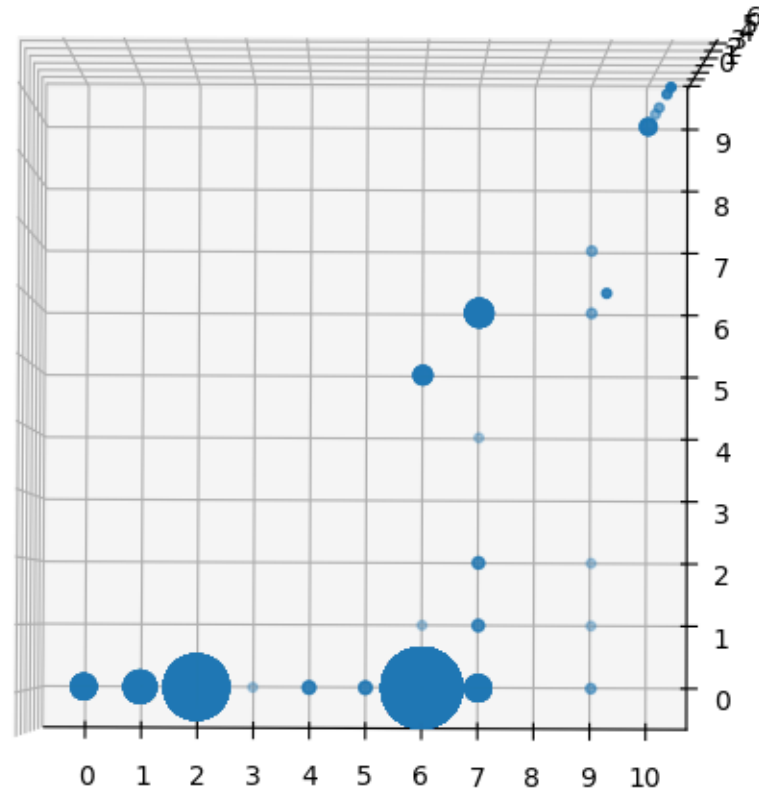
Razítka institucí společně 3D



Zdroj: vlastní výpočet a sběr dat

Razítka vyskytující se společně

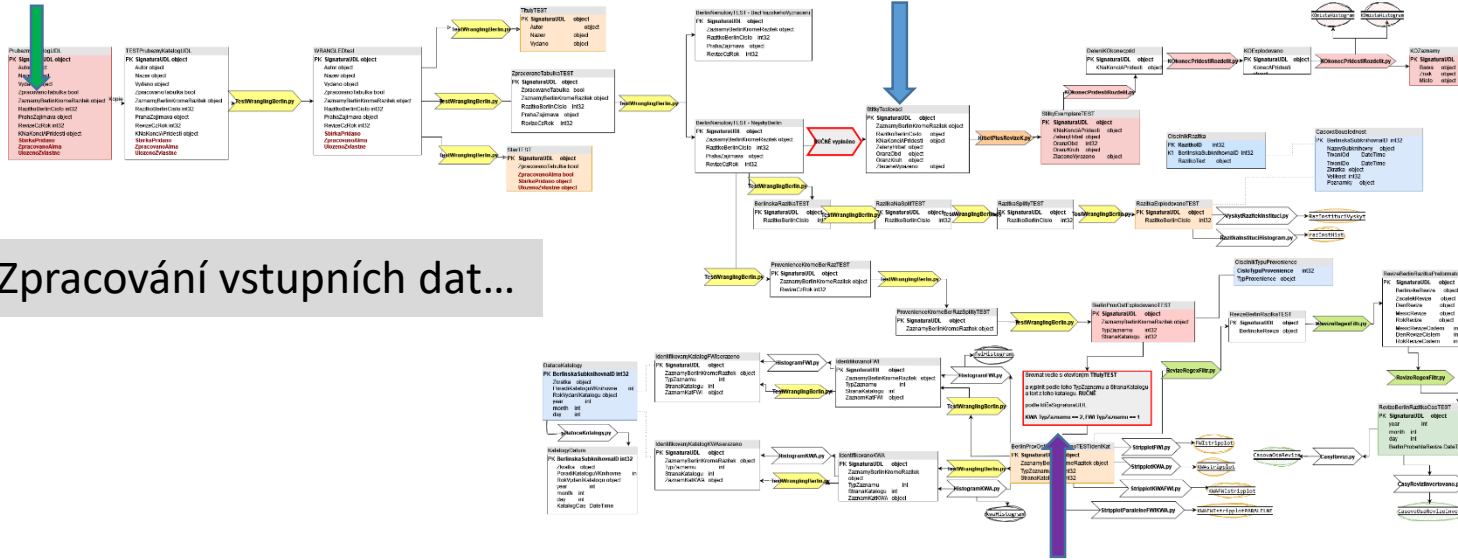
Razítka institucí společně 3D



Zdroj: vlastní výpočet a sběr dat; foto: archiv ÚVI

Výhody práce se skripty, nutné podmínky

- +
 - Přidat další záznamy a vše znovu zpracovat „na pár kliků“
 - Aktuální
- !
 - Držet si verze (př. Git)
 - Vědět, jaké parametry mám šanci přiřadit a identifikovat. Nastavit kategorie, oddělovače jednotlivých hodnot, číselné kódy, přehledy
- - Nevýhoda: pracnost prvotní výstavby „modelu“, námaha se vyplácí později



Zpracování vstupních dat...

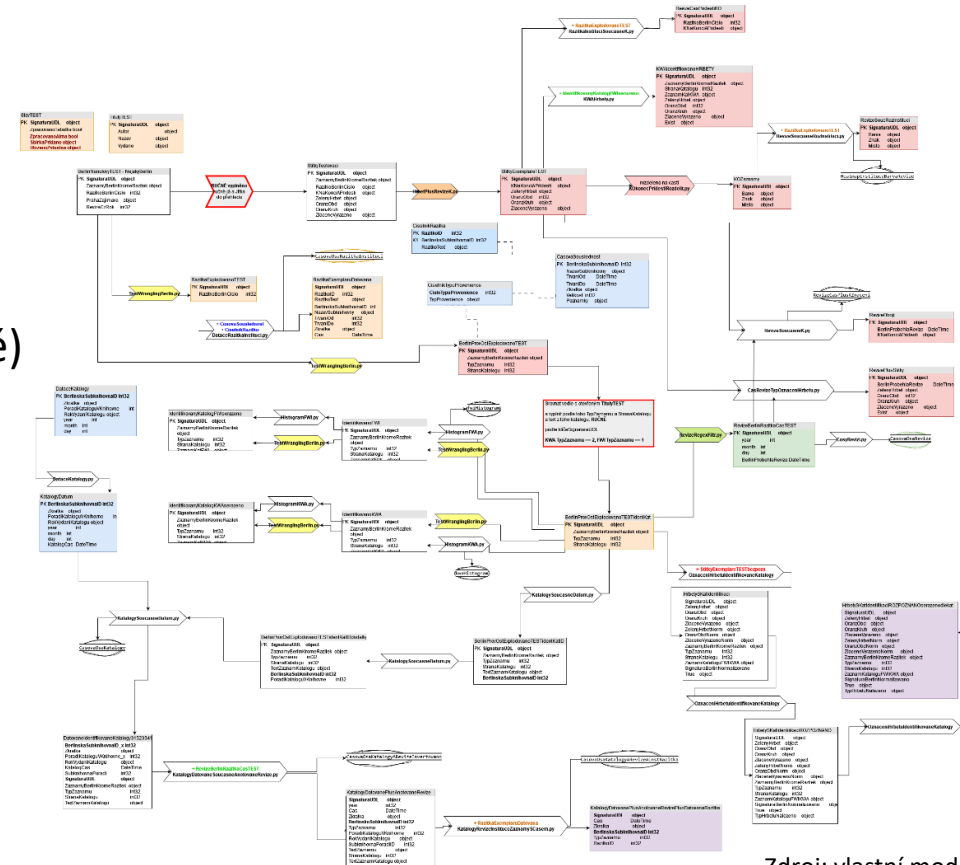
...a na ně navazuje kombinování dle klíčů a vykreslení grafů

Ruční kontrola s dochovanými katalogy

zatím 1/3 sbírky, proces replikovatelný (s novými daty vygeneruje znovu, stejně)

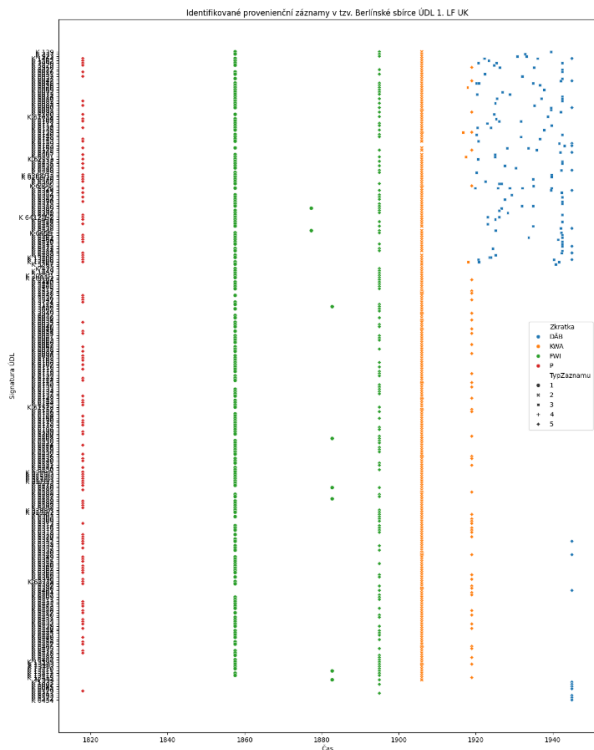
upřesnění podmínkami, vyexportovaná data, přiřazování hodnot dle nastavených klíčů

vypsat přehledové seznamy dle podmínek
vykreslit grafy, vizualizace vztahů, souvislostí

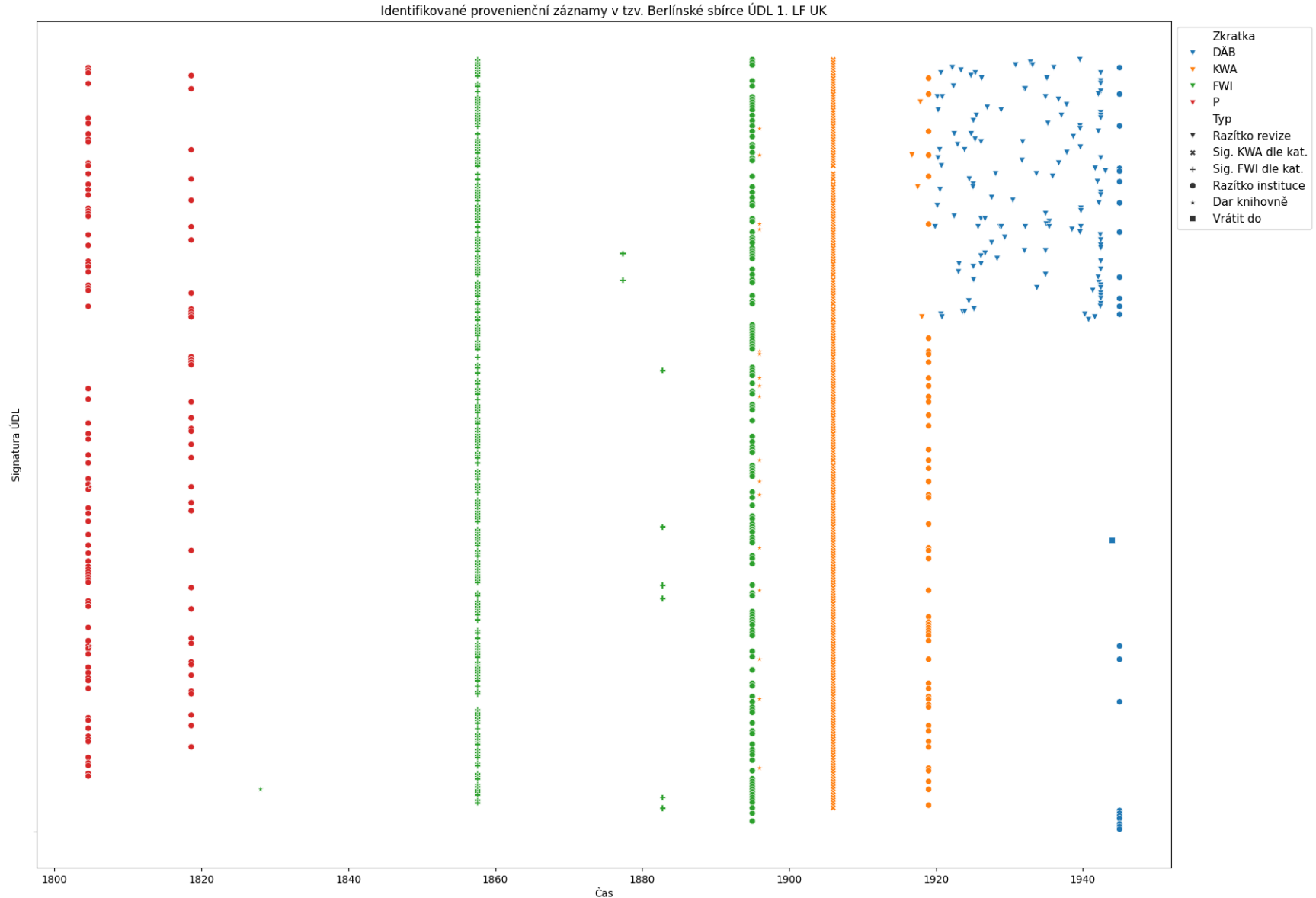


Zdroj: vlastní model

Příklad postupného upřesňování (přiřazení nového jména)

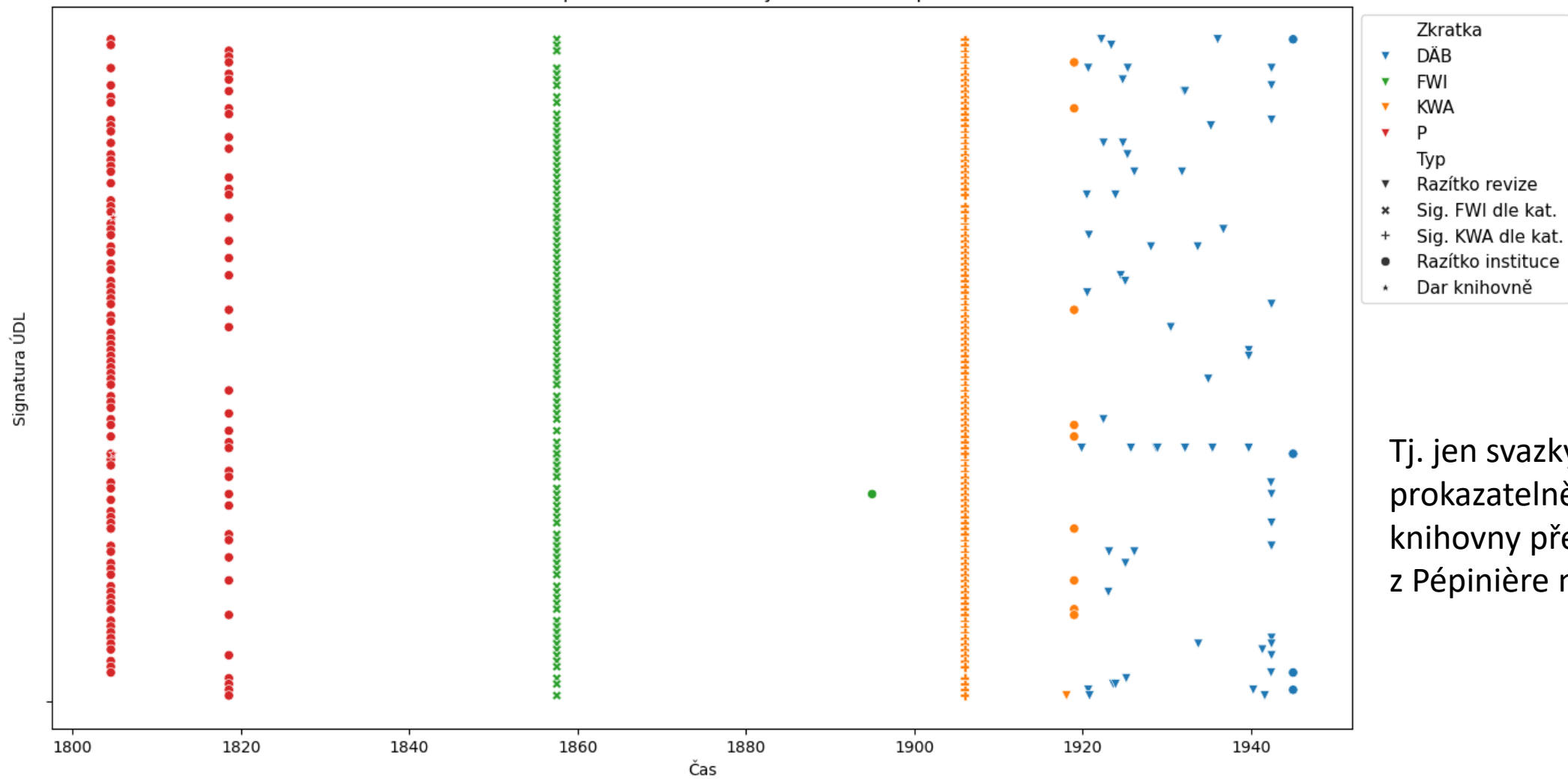


„Chirurgische Pépinière“
přejmenována 2. 8. 1804 na
„Medizinisch-chirurgische
Pépinière“ (*Ein geschichtlicher
Versuch*, s. 90), stačí ke kódům
razítek přiřadit další novou
proměnnou, překreslí se.



Zdroj: vlastní výpočet a sběr dat

Identifikované provenienční záznamy v knihách z Pépinière



Tj. jen svazky, které byly prokazatelně součástí knihovny před přejmenováním z Pépinière na FWI

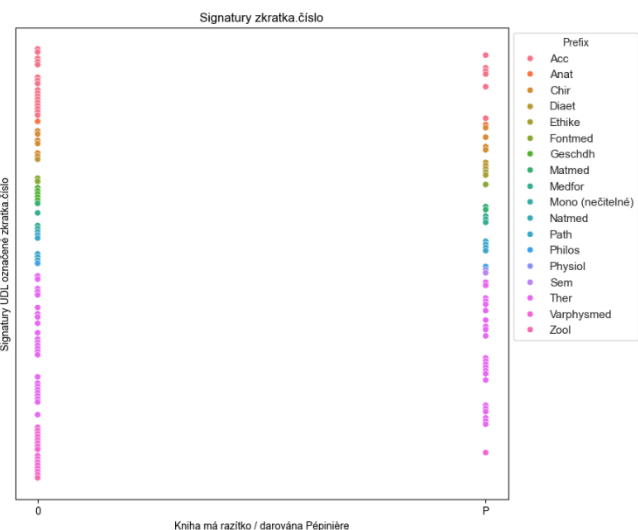
Zdroj: vlastní výpočet a sběr dat

Řada signatur **Zkratka.pořadové číslo**

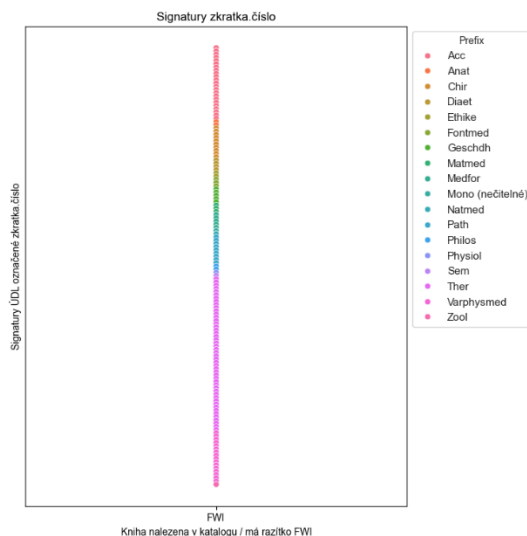
Nejasné, k čemu patří. Theußnerův (z. 1812) katalog Pépinière nemám
FWI má jiné signatury, mladší nebude.

1. Vypsat všechny neafiliované Zkratka.pořadové číslo – současný jasný identifikátor svazku
2. Vypsat všechny knihy se záznamem v katalogu FWI / všechny knihy s razítky Pépinière
3. Sloučit podle současného jasného identifikátoru svazku
4. Podívat se, s čím se překrývá

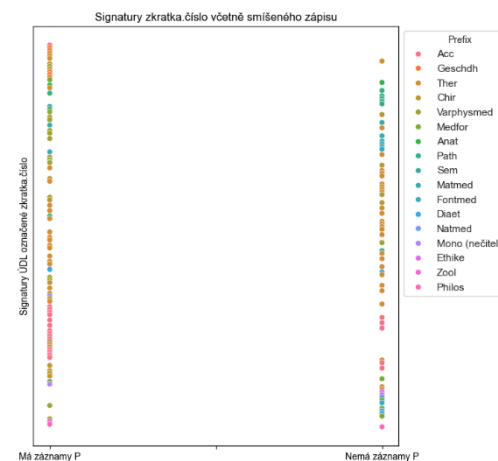
15 ks spojených signatur
Zkratka.číslo signatury FWI
s nimi totéž:



Pépinière záznamy
(nejpozději 1818)



Vůči katalogu či razítkům FWI
(1857; raz. teor. do 1895)



Nevystihuje P o nic lépe
> třetina jich ještě není
v knihovně 1818

Zdroj: vlastní výpočet a sběr dat

tedy ?
jako celek = podmnožina FWI

otestovat FWI z 2. stany:
vůči katalogu FWI (1857)
jen razítko nebude shoda



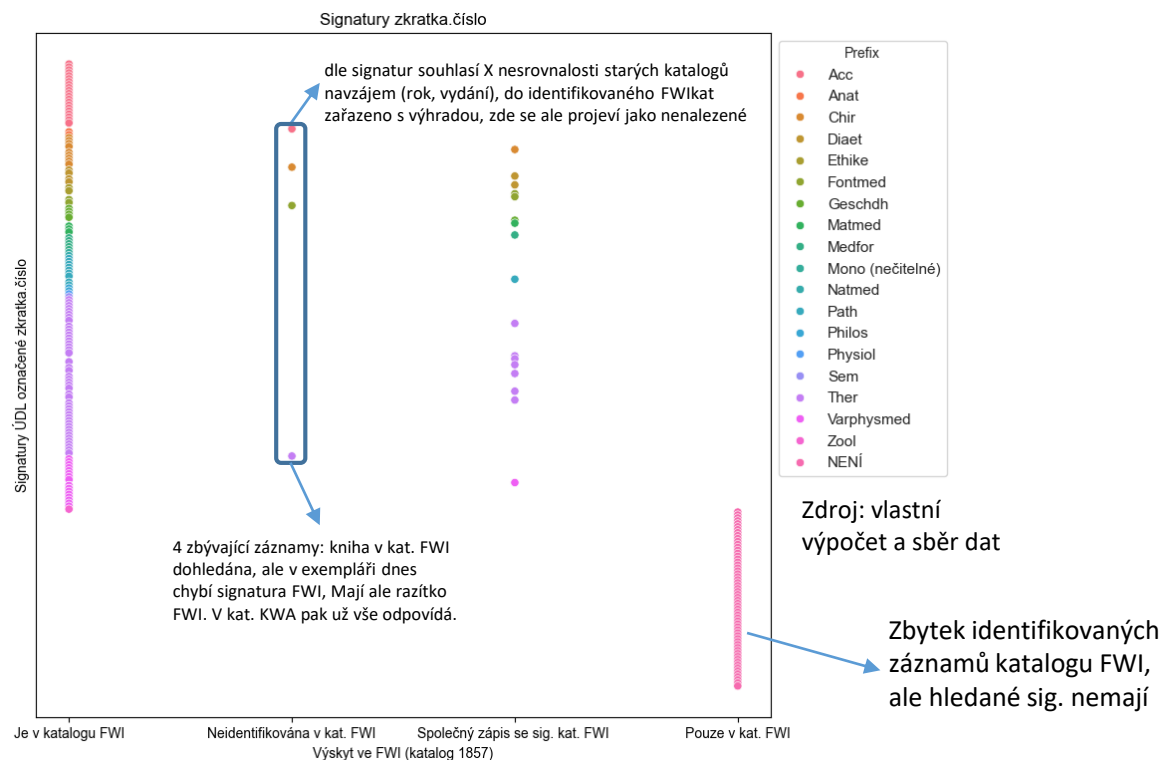
Cca „volný výběr“, který lépe odpovídá stavu v začátku FWI

Hypotéza:

Témata vyučovacích předmětů na **FWI** (přehled 1819)

Dokud knihy, které mám ve sbírce STT nezastaraly jako učební materiály

Poté pominul důvod jim připisovat sig., která byla od roku 1857 neaktuální



gende Ordnung der Vorlesungen ein:

Im 1ten Halbjahr: Physik und die gesammte Anatomie.

Im 2ten Halbjahr: Botanik, Naturgeschichte und Chemie; außerdem Enzyklopädie der Naturwissenschaften und der medizinischen Wissenschaften.

Im 3ten Halbjahr: die Anatomie praktisch und Physiologie.

Im 4ten Halbjahr: nochmals Physiologie und Chemie, Pharmazie und allgemeine Pathologie.

Im 5ten Halbjahr: allgemeine Semiotik, Materia medica, allgemeine Chirurgie, die Lehre von den primär-mechanischen Knochenkrankheiten und von den Bandagen.

Im 6ten Halbjahr: nochmals Materia medica, spezielle Pathologie und Chirurgie, Formulare und Therapie.

Im 7ten Halbjahr: theoretische und praktische Geburtshülfe, Therapie, Cursus operationum chirurgicarum, Augenheilkunde und Medicina forensis.

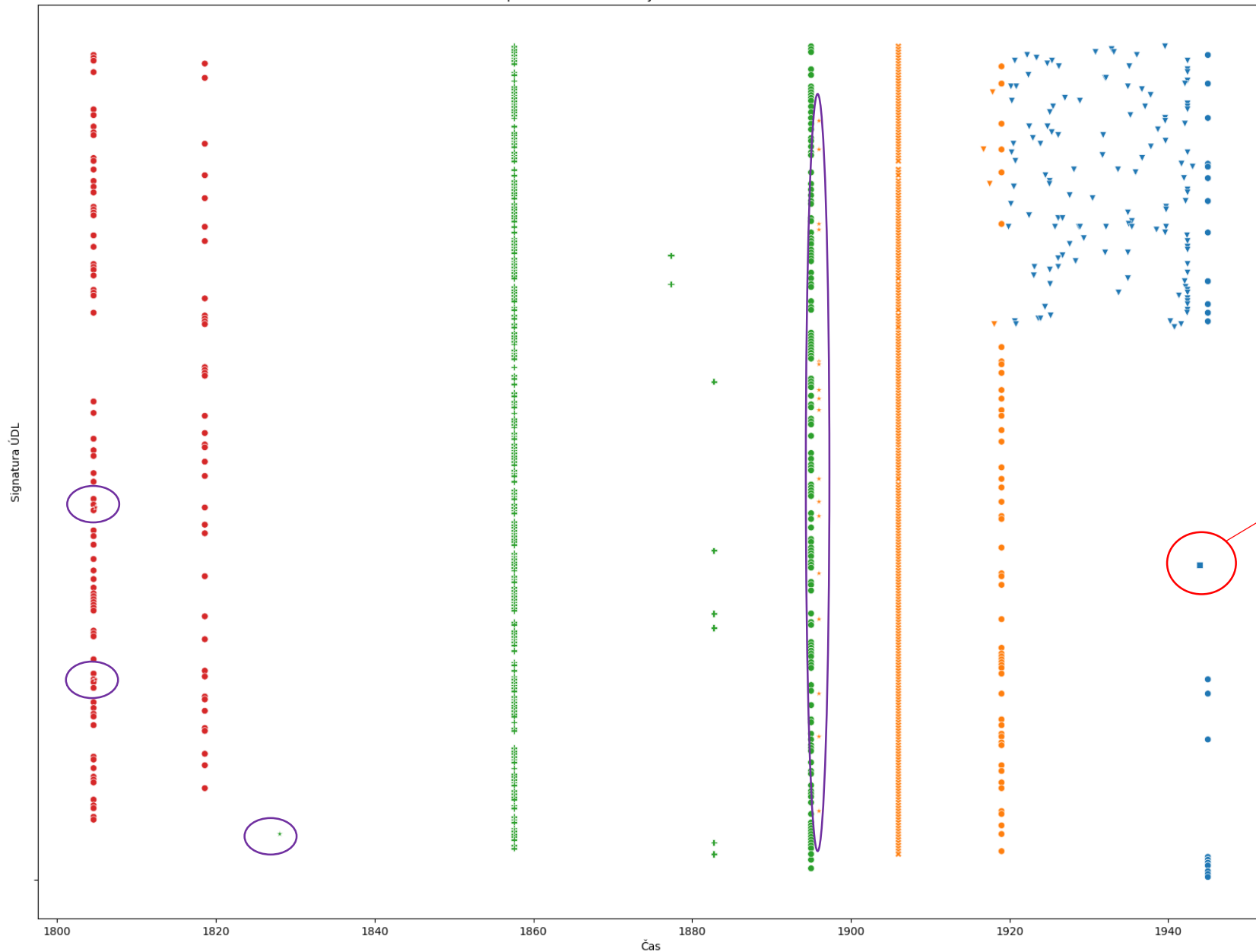
Im 8ten Halbjahr: Clinicum medicum, Polyclinicum, Clinicum chirurgicum et ophthalmiatricum. *).

FWI Stav výuky lékařství k roku 1819

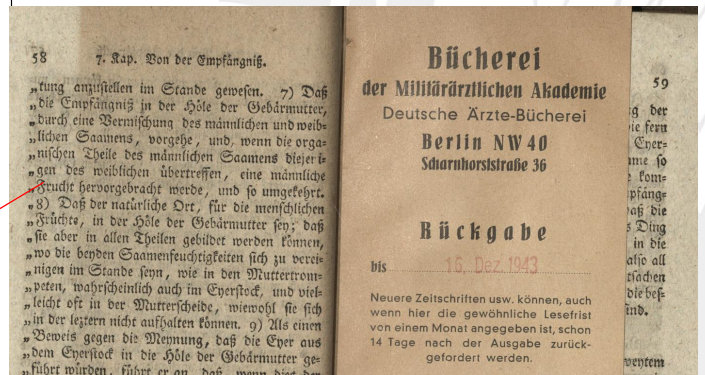
Preuß, J. D. E.: *Das Königlich Preußische medizinisch-chirurgische Friedrich-Wilhelm's-Institut (ursprünglich chirurgische Pépinière) zu Berlin. Ein geschichtlicher Versuch zum 25ten Stiftungstage desselben dem 2ten August 1819.* Berlin: Gedruckt in der Ungerschen Buchdruckerei, 1819, s. 59n.

Celkem 10 000 svazků

Katalog FWI 1857 už má 27 000 svazků



- Zkratka
- ▼ DAB
 - ▼ KWA
 - ▼ FWI
 - ▼ P
 - Typ
 - ▼ Razítko revize
 - × Sig. KWA dle kat.
 - + Sig. FWI dle kat.
 - Razítko instituce
 - Dar knihovně
 - Vrátit do



K 6229

* jmenovité dary

Zdroj: vlastní výpočet a sběr dat

„tung anzustellen im Stande gewesen. 7) Daß
 „die Empfängniß in der Höle der Gebärmutter,
 „durch eine Vermischung des männlichen und weib-
 „lichen Saamens, vorgehe, und, wenn die orga-
 „nischen Theile des männlichen Saamens diejeri-
 „gen des weiblichen übertreffen, eine männliche
 „Frucht hervorgebracht werde, und so umgekehrt.
 „8) Daß der natürliche Ort, für die menschlichen
 „Früchte, in der Höle der Gebärmutter sey; daß
 „sie aber in allen Theilen gebildet werden können,
 „wo die beyden Saamenfeuchtigkeiten sich zu verei-
 „nigen im Stande seyn, wie in den Muttertrom-
 „peten, wahrscheinlich auch im Eyerstock, und viel-
 „leicht oft in der Mutterscheide, wiewohl sie sich
 „in der letztern nicht aufhalten können. 9) Als einen
 „Beweis gegen die Meinung, daß die Eyer aus
 „dem Eyerstock in die Höle der Gebärmutter ge-
 „führt würden, führt er an, daß, wenn dies der

Bücherei

der Militärärztlichen Akademie

Deutsche Ärzte-Bücherei

Berlin NW 40

Sdiarnhorststraße 36

Rückgabe

bis

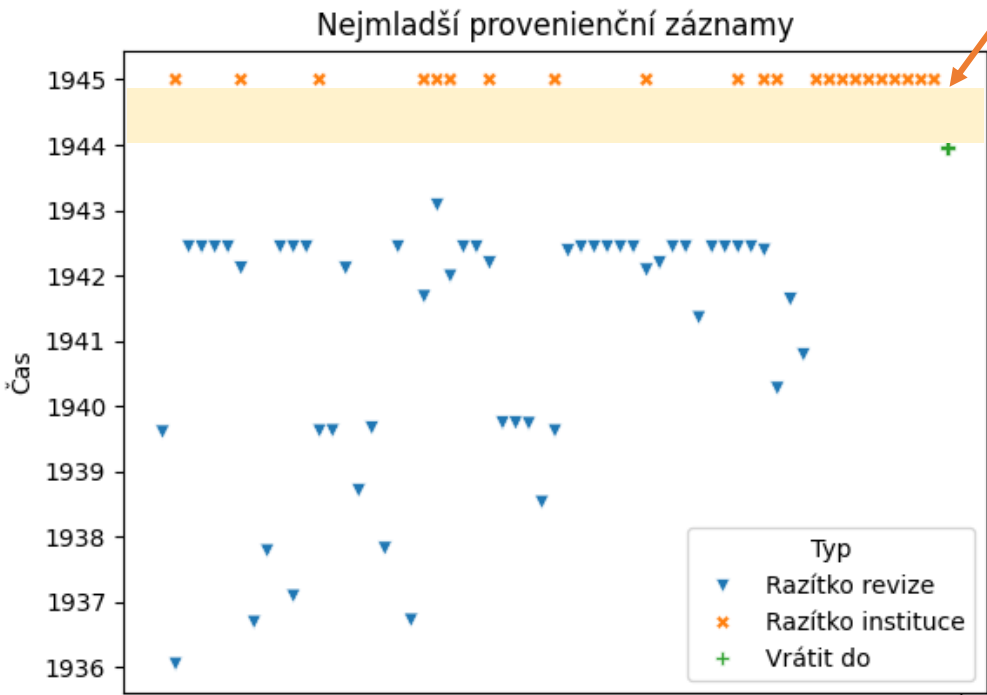
16. Dez. 1943

Neuere Zeitschriften usw. können, auch wenn hier die gewöhnliche Lesefrist von einem Monat angegeben ist, schon 14 Tage nach der Ausgabe zurückgefordert werden.

g der
 die fern
 Eyer-
 me so
 e kom-
 pfäng-
 daß die
 s Ding
 in die
 also all
 tsachen
 die bef-
 ind.

ventem

- Jinak podle razítek a signatur nezajímavě působící svazek.
- Nejzazší datum razítek DÄB nastaveno na 1945, x jen konstrukt teoreticky nejpozdějšího data (zánik celé instituce). Kdykoliv posunout.
- Ostatní – dosud zpracovaná – razítka revizí jsou starší než „vratte do“.
- Pak máme až zápisy přírůstkového seznamu ÚDL 1954 – kdy jsou už součástí fondu (a české revize fondu ÚDL) jako „darem“, což ovšem označuje i svazky, které byly součástí fakultní knihovny i před válkou.
- Víc nám knihy samy nepoví, ale revizní datum z Berlína vidíme před sebou kontinuálně, knihy „nevisí před koncem války ve vzduchu“.



Signatura ÚDL Zdroj: vlastní výpočet a sběr dat

„Bekannt ist ferner durch Augenzeugen, daß im Sommer 1945 die **nicht ausgelagerten Bestände** der reichen Deutschen Ärzte-Bücherei – 1942 zählte sie 133.000 Bände und 120.000 Dissertationen und ging auf die Pépinière des 18. Jahrhunderts zurück – aus dem stark beschädigten Gebäude **durch sowjetisches Militär abtransportiert worden sind.**“ *Handbuch*, s. 33. (s pozn.: „Ich beziehe mich auf eine Information des **Berliner Mediziners Prof. Dr. Wolfram Körner.**“)

t. ante quem

- Hranici opuštění Berlína **lze posunout** v závislosti na zjištěných faktech (A) evakuace před koncem války, viz *Knihy znovu nalezené a Handbuch der historischen Buchbestände* s. 30, nebo B) zabrání zbytku knihovny po konci války, viz *Handbuch* s. 33 a nálezy v Gruzii, táž razítka a signatury FWI a KWA).
- Stačí změnit 1 záznam v tabulce datací razítek a ve správném pořadí pustit správné skripty.
- Evakuace počátkem 1945 neodporuje termínu vrácení K 6229. Eventualitu zabrání zbytku neevakuovaného fondu z dat nelze prokázat ani vyloučit. Předpokládat evakuaci do Českého Dubu lze, ale jen na základě analogií.

„nach 1941 aber mit Theilen ihres Hauptbestandes, auf der Flucht aus der Hauptstadt zu verschiedenen, über das ganze damalige Reichsgebiet verstreuten Bergungsorten befanden. Millionen von Bändern wurden durch ganz Deutschland bewegt. Die Preußische Staatsbibliothek nahm **Anfang 1945 eine nahezu vollständig durchgeführte Gesamtverlagerung** (...) an 30 Bergungsorte vor, die später zu drei Besatzungszonen, zur **Tschechoslowakei** und zu Polen gehörten.“ *ibid.* s. 30.

QED

- **Vyplatí se, ale...**
- Ideálně, když už data máte sesbíraná, např. zpracované záznamy v katalogu a z něj vyexportované údaje s kódovanými kategoriemi
- Nový vhled, **přidanou hodnotu má**. Pomáhá hledat interpretace
 - Neidentifikované signatury – FWI
 - Hřbety většinou označené signaturou KWA (\pm spadlé štítky, zůstala jen starší spodní vrstva později plošně ve fondu přelepená KWA)
 - Zpřesnění posledních záznamů z Berlína
 - Mapování individuálních darů
 - Vymezení celku „berlínské sbírky“

- **Data:** vlastní sběr dat provenienčních záznamů sbírky ÚDL 1. LF UK, archiv ÚDL a ÚVI.
- Preuß, J. D. E.: *Das Königlich Preußische medizinisch-chirurgische Friedrich-Wilhelm's-Institut (ursprünglich chirurgische Pépinière) zu Berlin. Ein geschichtlicher Versuch zum 25ten Stiftungstage desselben dem 2ten August 1819.* Berlin: Gedruckt in der Ungerschen Buchdruckerei, 1819.
- *Katalog der Bibliothek des Koeniglichen medizinisch-chirurgischen Friedrich-Wilhelms-Instituts.* Berlin: Druck der Gebr. Unger'schen Hofbuchdruckerei, 1857.
- *Verzeichnis der Büchersammlung der Kaiser Wilhelms-Akademie für das militärärztliche Bildungswesen.* (Dritte Ausgabe). Berlin: Verlag von August Hirschwald. Unter den Linden 68, 1906.
- Schätze aus der Berliner Ärzte-Bibliothek. UB Bochum. dostupné na: <https://www.ub.ruhr-uni-bochum.de/DigiBib/Aktuelles/Schaetze.html> staženo 26. 11. 2023.
- „Berliner Ärzte-Bibliothek. Kurztitel der 2004 von Georgien zurückgegebenen ca. 900 Bände“. dostupné na: <https://www.ub.ruhr-uni-bochum.de/projekte/dateien/aerzte-bibliothek.pdf> staženo 26. 11. 2023.
- *Handbuch der historischen Buchbestände. Berlin. Theil 1.* Fabian B. (ed.). Hildesheim: Olms-Weidmann, 1995.
- Vollbeding, J. C.: *Neuer gemeinnützlicher Briefsteller für das bürgerliche Geschäftsleben.* Berlin: C. F. Amelang, 1825.
- *Neuester Correspondent und Geschaeftrath.* Wien: In der Carl Haas'schen Buchhandlung, 1829.
- *Jahrbuch der Deutschen Bibliotheken.* Verein Deutscher Bibliothekare. 1. Jahrgang. Leipzig: Otto Harrassowitz, 1902.
- **Software:** vlastní skripty pro tvorbu modelu proveniencí a jejich vizualizace (použit Python verze 3.11.2, Pandas 1.5.3, Matplotlib 3.8.0, Seaborn 0.13.0... a online nástroj Regular Expressions 101).

regular expressions 101

social donate info

SAVE & SHARE

- Save new Re... `ctrl+s`
- Update Re... `ctrl+u+s`
- Delete Regex
- Add to Community L...

FLAVOR

- PCRE2 (PHP >=7.3)
- PCRE (PHP <7.3)
- ECMAScript (JavaScri...)
- Python** ✓
- Golang
- Java 8
- .NET (C#)
- Rust

FUNCTION

- Match** ✓
- Substitution
- List

REGULAR EXPRESSION `1 match (9 steps, 1.0ms) v1`

```

r"\d\d?[.]?[ ]?[a-zA-Z]{3}[.]?[ ]?\d{4}"

```

TEST STRING

6Sep1940

12. NOV. 1935

5 AUG. 1924

6SEP1940

6 .SEP. 1940

22SEP1939

12. nov. 1935

5 aUg. 1924

22seP1939

EXPLANATION

in the list below `[a-zA-Z]`

- `{3}` matches the previous token exactly 3 times
- `a-z` matches a single character in the range between a (index 97) and z (index 122) (case sensitive)
- `A-Z` matches a single character in the range between A (index 65)

MATCH INFORMATION

Match 1

QUICK REFERENCE

Search referen

- All Tokens
- Common T

PLEASE SUPPORT REGEX101



<https://xkcd.com/208/>