**Library Carpentry (URL:** [**https://librarycarpentry.org/**](https://librarycarpentry.org/)**)**

Library Carpentry razvija radionice koje su namijenjene za osobe koje rade u knjižničnim i informacijskim ulogama. Nastoje osposobiti druge za učinkovite, djelotvorne i ponovljive prakse podataka i softvera. Vještine kojima ih nastoje osposobiti su:

* Savladati termine relevantne za razvoj softvera i obradu podataka kako bi ih mogli primjenjivati unutar knjižnica
* Identificirati i primjenjivati najbolju praksu za strukturiranje podataka
* Transformirati i mapirati podatke iz jednog oblika u drugi
* Učinkovitije raditi/surađivati s istraživačima, IT kolegama i kolegama iz njihove struke
* Automatizirati podatke koji su skloni ponavljanju

Lekcije uvode pojmove, fraze i koncepte u razvoju softvera i znanosti o podacima, kako najbolje raditi s strukturama podataka i koristiti regularne izraze u pretraživanjima. Vježbe pokrivaju brojanje i otkrivanje podataka. Osim toga, pokrivaju i rad s OpenRefineom kako bi transformirali i očistili podatke te prednosti zajedničkog rada putem Git / GitHub-a i pomoću kontrole verzija za praćenje rada.

Kako bi se zadovoljile potrebe korisnika na fleksibilan način, knjižničari se koriste kodiranjem kao sredstvom kojim preuzimaju kontrolu nad praksom i tako osnažuju sebe i svoju organizaciju. Važno je poznavanje UNIX-a, GIT-a, SQL-a, Python-a, podataka i arhivistike. Kako bi se olakšao posao pretraživanjai organiziranja podataka, važno je naučitiUNIX Shell. Prilikom kopiranja često mogu izazvati probleme velike količine podataka ili digitalne datoteka raštrkane po računalu ili udaljenom poslužitelju prilikom kopiranja, premještanja, preimenovanja, prebrojavanja, pretraživanja ili ručnog procesiranja tih datoteka.

Pojam ''Ljuska'' (koja se ponekad naziva "Unix ljuska", zbog operativnog sustava u kojem je prvi put razvijena) predstavlja program koji omogućuje interakciju s računalom pomoću tipkanih tekstualnih naredbi. To je primarno sučelje koje se koristi na Linux i Unix sustavima, kao što je MacOS, i može se instalirati opcionalno na druge operativne sustave kao što je Windows. To je primjer "sučelja komandne linije", gdje se upute daju računalu upisivanjem naredbi, zatim računalo odgovara izvršavanjem zadatka ili generiranjem izlaza. Ljuska je vrlo produktivan program koji se može upotrijebiti za automatiziranje svojega rad.

GitHub je razvojna platforma za kontrolu različitih verzija podataka. To je besplatan alat otvorenog koda kojim može pregledati kod, upravljati projektima, izraditi softver i koristiti za bilježenje svih promjena koje su napravljene u grupi određenih računalnih datoteka (koje se kratko nazivaju "git repozitorij" ili "repo"). GitHub predstavlja web-lokaciju za udaljeno pohranjivanje i dijeljenje Git repozitorija.

OpenRefine je moćan alat koji služi za rad s neurednim podacima. Dakle, može se koristiti za standardizaciju i čišćenje podataka u datoteci, za pretvaranje jednog formata u drugi te za proširivanje web-usluga i vanjskih podataka.

SQL **(**Structured Query Language) je moćan programski jezik koji služi za izradu, traženje, ažuriranje i brisanje podataka iz relacijskih baza podataka. SQL je optimiziran za rukovanje velikim količinama podataka i standardiziran je preko standarda ANSI i ISO. On može biti vrlo koristan knjižničarima za izravno slanje upita unutar baza podataka bez ograničenja.

Web scraping **(**''web struganje'') je tehnika za vađenje informacija s web stranica. On nam omogućuje da pretvorimo ne-tablične ili slabo strukturirane podatke u upotrebljiv i strukturiran format, kao što je .csv datoteka ili proračunska tablica. ''Struganje'' nam može pomoći da dobijemo one podatke koji se na prvu čine nedostupnima na webu. Posjeduje li neka informacijska ustanova vlastitu stranicu, pomoću Web scrapinga može obrađivati web-stranicu te na web-stranici dohvaćati i vaditi podatke iz nje.

**Data Carpentry (URL:** [**https://datacarpentry.org/**](https://datacarpentry.org/)**)**

Data Carpentry razvija i podučava radionice o temeljnim vještinama podataka potrebnim za provođenje istraživanja. Nastoje pružiti istraživačima visokokvalitetnu obuku koja pokriva cijeli životni ciklus istraživanja temeljenih na podacima.

Njihove lekcije usmjerene su na podučavanje sljedećih polja znanja:

* Ekologija
* Genomika
* Društvene znanosti
* Geoprostorni podaci

1. Radionica za Ekologiju koristi tabelarni skup podataka o ekologiji iz baze podataka projekta Projekt portala i predaje čišćenje podataka, upravljanje, analizu i vizualizaciju. Tijekom radionice koriste jedan skup podataka kako bi modelirali tijek rada za upravljanje podacima i analizu koji bi istraživač koristio. Radionica Ekologija može se podučavati koristeći R ili Python kao osnovni jezik.
2. Radionica za genomiku usmjerena je na rad s podacima o genomici te upravljanju podacima i analizama za istraživanje genomike, uključujući najbolje prakse za organizaciju bioinformatskih projekata i podataka, korištenje alata za naredbe, povezivanje i korištenje računalstva u oblaku...
3. Radionica za društvene znanosti koristi skup podataka iz SAFI Teaching baze podataka i poučava čišćenje podataka, upravljanje, analizu i vizualizaciju. Radionica društvenih nauka može se podučiti koristeći R kao osnovni jezik.
4. Radionice za geoprostorne podatke usredotočena je se na rad s geoprostornim podacima odnosno upravljanje i razumijevanje formata prostornih podataka, razumijevanje koordinatnih referentnih sustava i rad s rasterskim i vektorskim podacima u R-u za analizu i vizualizaciju.

Python je programski jezik opće namjene koji podržava brz razvoj aplikacija za analizu podataka. Riječ "Python" koristi se za upućivanje na programski jezik i alat koji izvršava skripte napisane na Python jeziku. Po automatskoj memorijskoj alokaciji, Python je sličan programskim jezicima kao što su Perl, Ruby, Smalltalk itd. On pruža konstrukte koji omogućuju jasno programiranje na malim i velikim razmjerima. Pomoću njega se izrađuju programi koji olakšavaju i ubrzavaju uobičajene funkcije.

**Software Carpentry (URL:** [**https://software-carpentry.org/**](https://software-carpentry.org/)**)**

Software Carpentry (1998-.) stranica podučava istraživače računalnim vještinama koje su im potrebne da bi se postigla što veća učinkovitost u što manje vremena i uz manje boli. Nagrađivani i educirani instruktori podučavaju istraživače o trima osnovnim temama: Unix ljuska, programski jezik (Python ili R) i kontrola verzija s Git-om. Nude se i dodatne korisne lekcije programiranja koje obuhvaćaju programiranje pomoću GAP-a, MATLAB-a SQL-a…

MATLAB vlasnički programski jezik visoke razine koji je razvio MathWorks i predstavlja interaktivnu okolinu za numeričko i matrično računanje. On omogućuje nam da vidimo kako različiti algoritmi rade s našim podacima, manipulaciju matricama, crtanje funkcija i podataka, implementaciju algoritama, stvaranje korisničkih sučelja i povezivanje s programima pisanim na drugim jezicima, uključujući C, C ++, C #, Java, Fortran i Python.

GAP je sustav za računsku diskretnu alegebru s posebnim naglaskom na teoriji računskih grupa. GAP pruža programski jezik, biblioteku tisuća funkcija koje implementiraju algebarske algoritme napisane u GAP jeziku, kao i velike biblioteke podataka algebarskih objekata. se koristi u istraživanju i poučavanju za proučavanje grupa i njihovih prikaza, prstena, vektorskih prostora, algebri, kombinatornih struktura i još mnogo toga.

Mercurial je besplatan alat za upravljanje kontrolom izvora. Nudi vam moć da učinkovito rukujete projektima bilo koje veličine dok koristite intuitivno sučelje. Jednostavan je za korištenje i teško ga je razbiti, što ga čini idealnim za svakoga tko radi s verzioniranim datotekama.