NAČELA METODOLOGIJE ZNANSTVENOG RADA (298870; 60246)

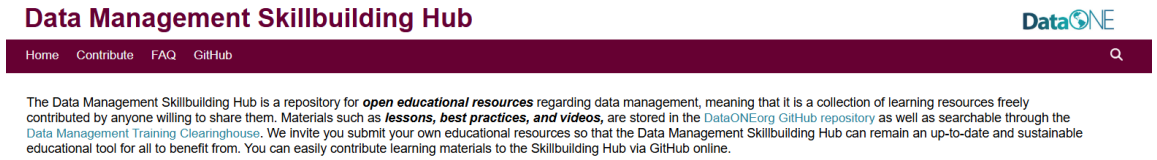
Akademska godina 2026/2027
Nositelji: prof.dr.sc. Marko Miliša; **izv.prof.dr.sc. Sunčica Bosak**

UVOD U PODATKOVNU PISMENOST U (EKOLOŠKIM) ISTRAŽIVANJIMA – „DATA LITERACY”

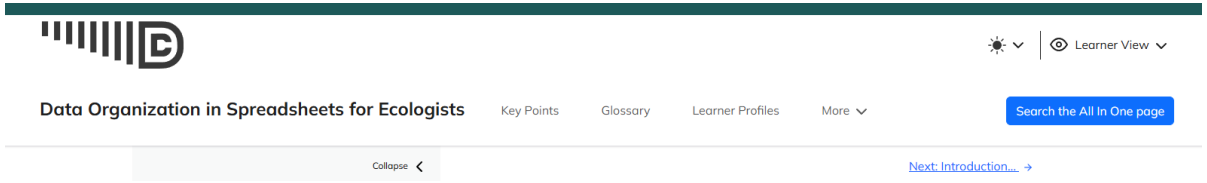
- Što je podatkovna pismenost?
- Što su ekološki podaci?
- Od istraživačkog pitanja do podataka
- Kontekst, metapodaci i organizacija
- Kvaliteta i ograničenja podataka
- Čitanje prikaza i interpretacija

Literatura:

https://dataoneorg.github.io/Education/



https://datacarpentry.github.io/spreadsheet-ecology-lesson/instructor/index.html







Podatkovna pismenost u ekologiji (biologiji) znači razumjeti put od opažanja u prirodi do znanstvene tvrdnje - i prepoznati odluke, pretpostavke i ograničenja na tom putu.

ŠTO JE PODATKOVNA PISMENOST? „DATA LITERACY”

Dobili ste tablicu s rezultatima terenskog istraživanja. Što morate znati prije nego što počnete računati prosjeke ili izrađivati grafove?

Lokalitet A	Lokalitet B
40 vrsta	25 vrsta

Gdje je zabilježeno više vrsta?

ili

Koji je lokalitet bogatiji vrstama?

Klorofil a = 3,2

U kojoj mjernoj jedinici?
U kojem uzorku i na kojoj dubini?
Kada i gdje je uzorak prikupljen?
Kojom je metodom određena koncentracija?
Je li riječ o pojedinačnom mjerenju ili prosjeku?

Lokalitet A	Lokalitet B
40 zabilježenih vrsta	25 zabilježenih vrsta
10 terenskih istraživanja	1 terensko istraživanje
Travanj – rujan	lipanj

Da li bi ste isto zaključile s obzirom na nove informacije?

Brojke se nisu promijenile. Dodali smo informacije o tome kako su nastale

Što provjeriti: način opažanja, napor uzorkovanja, obuhvaćene skupine organizama, vrijeme i prostorni obuhvat

METAPODACI

opis zašto, tko, što, kada, gdje i kako je prikupljao podatke.

Ta pitanja su osnova dokumentiranja podataka

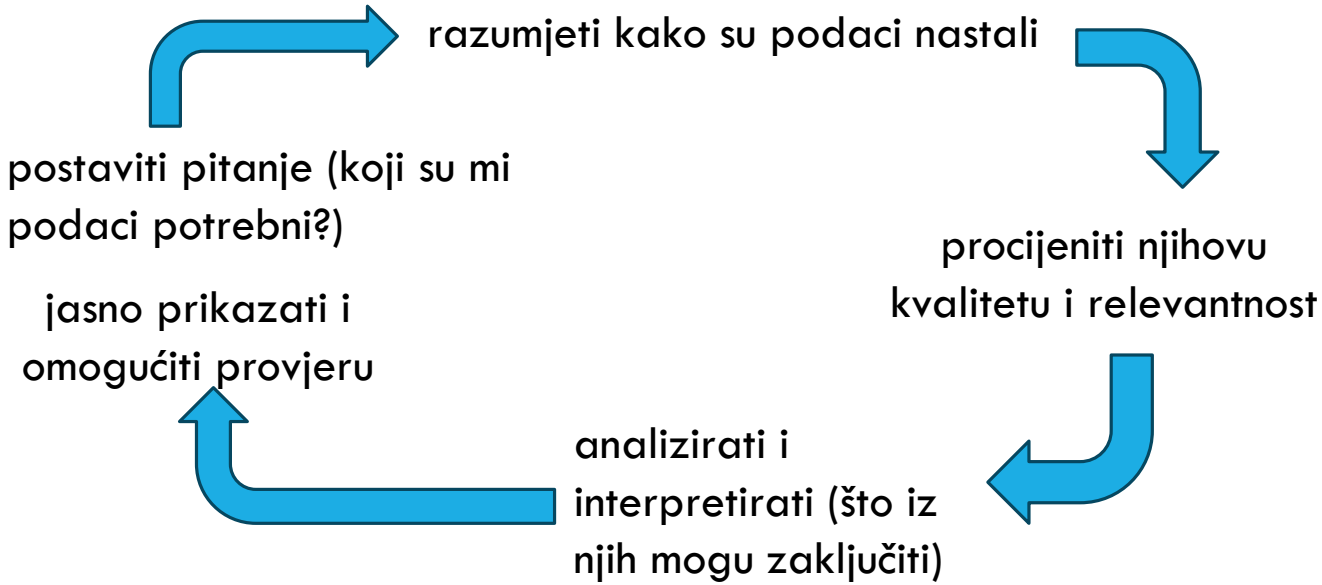
Metapodaci su strukturirane informacije koje opisuju, objašnjavaju, lociraju ili daju kontekst drugim podacima, a često se definiraju jednostavno kao „podaci o podacima

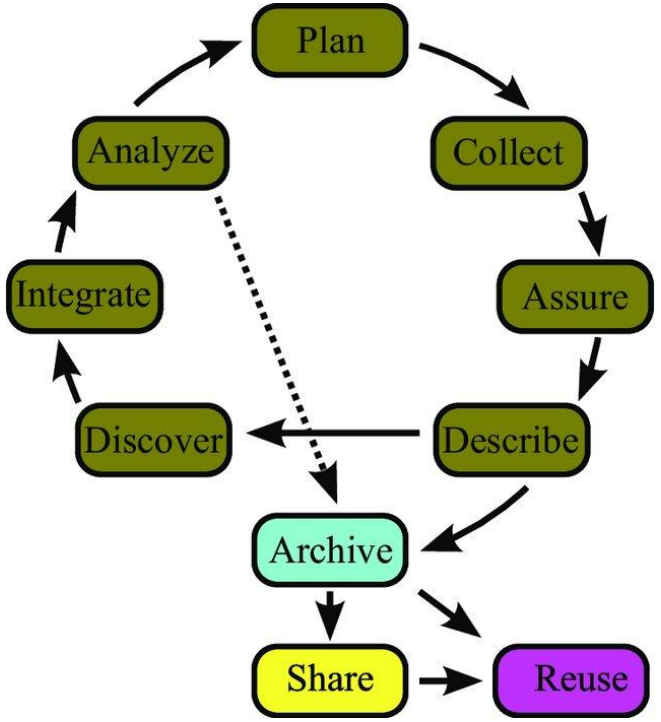
Zamislite da ovu tablicu dobijete za diplomski rad, bez kontakta s osobom koja je prikupila podatke. Biste li imale dovoljno informacija za rad?”

Lokalitet A	Lokalitet B
40 vrsta	25 vrsta

To je podatkovna pismenost (data literacy).

Podatkovna pismenost sposobnost je razumijevanja podrijetla i značenja podataka, procjenjivanja njihove prikladnosti te odgovornog interpretiranja i komuniciranja zaključaka.





www.dataone.org/data-life-cycle

ŠTO SVE MOŽE BITI ISTRAŽIVAČKI PODATAK??

Različiti oblici ekoloških podataka

Primjer	Što trebamo dodatno znati?
Broj jedinki neke vrste	Površinu, volumen ili trajanje uzorkovanja
Fotografija mikroskopskog preparata	Mjerilo, oznaku uzorka, način snimanja
Nukleotidna sekvenca	Podrijetlo uzorka i postupak dobivanja sekvence
Zapis prisutnosti vrste	Vrijeme, lokaciju i metodu opažanja
Vremenski niz temperature	Jedinicu, učestalost mjerenja i podatke o senzoru
Terenska bilješka	Autora, datum i okolnosti opažanja

ŠTO PREDSTAVLJA JEDNA TOČKA ILI JEDAN REDAK?

sample_id	lokalitet	temperatura_C	klorofil_a_ug_L
U001	A	18,4	1,2
U002	A	18,7	1,4
U003	B	20,1	2,6

Što se mijenja ako u istom uzorku određujemo brojnost 50 vrsta?

Istraživačko pitanje određuje koje podatke trebamo

Razlikuje li se koncentracija klorofila a između dvaju obalnih lokaliteta tijekom ljeta?

- koje lokalitete uspoređujemo;
- kada i koliko često uzorkujemo;
- na kojoj dubini;
- koliko uzoraka prikupljamo;
- kojom metodom određujemo klorofil a;
- koje dodatne okolišne podatke bilježimo.

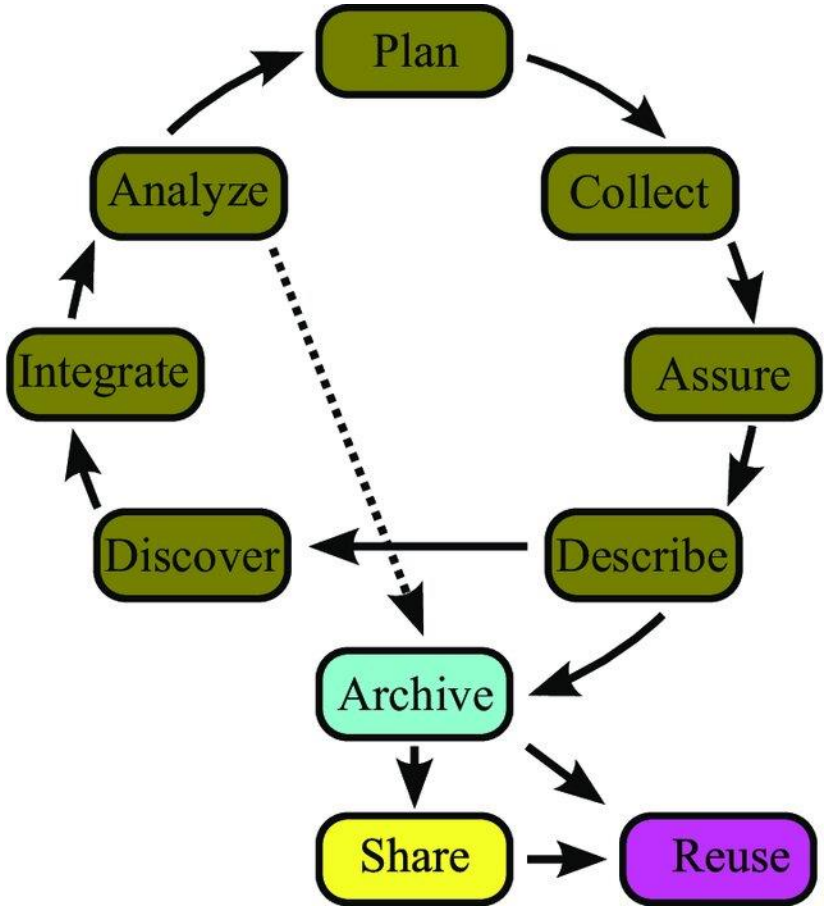
Povezivanje pitanja s relevantnim podacima važan je dio podatkovne pismenosti u ekologiji

Hitchcock, C., Sullivan, J. and O'Donnell, K. (2021) 'Cultivating Bioliteracy, Biodiscovery, Data Literacy, and Ecological Monitoring in Undergraduate Courses with iNaturalist', Citizen Science: Theory and Practice, 6(1), p. 26. Available at: https://doi.org/10.5334/cstp.439.

Ekološki fenomen → način opažanja ili mjerenja → zabilježena varijabla → interpretacija

Želimo istražiti fitoplanktonsku zajednicu. Hoćemo li bilježiti sastav vrsta, brojnost stanica, biovolumen ili koncentraciju klorofila a? Što točno želimo saznati?

Planiranje je i početak životnog ciklusa podataka.



U kojoj fazi trebamo početi pisati što znače oznake u tablici?

www.dataone.org/data-life-cycle

METAPODACI

Metapodaci su informacije koje opisuju podatke, njihovo podrijetlo, strukturu i kontekst.

- oznaka uzorka;
- datum, vrijeme i lokacija;
- dubina uzorkovanja;
- metoda prikupljanja i analize;
- naziv varijabli i mjerne jedinice;
- značenje šifri i oznaka nedostajućih vrijednosti;
- napomene o kvaliteti.

opisivanje podataka kroz pitanja zašto, tko, što, kada, gdje i kako

dva praktična dokumenta:

README - objašnjava skup podataka i njegov kontekst.

Rječnik podataka - objašnjava pojedine varijable, jedinice i dopuštene vrijednosti.

Izrada rječnika podataka jedna je od preporuka za pouzdanu organizaciju tabličnih podataka.

	A	B	C	D
1	name	plot_name	group	description
2	mouse	Mouse	demographic	Animal identifier
3	sex	Sex	demographic	Male (M) or Female (F)
4	sac_date	Date of sac	demographic	Date mouse was sacrificed
5	partial_inflation	Partial inflation	clinical	Indicates if mouse showed partial pancreatic inflation
6	coat_color	Coat color	demographic	Coat color, by visual inspection
7	crumblers	Crumblers	clinical	Indicates if mouse stored food in their bedding
8	diet_days	Days on diet	clinical	Number of days on high–fat diet

Broman, K. W., & Woo, K. H. (2018). Data organization in spreadsheets. *The American Statistician*, *72*_(1), 2–10. https://doi.org/10.1080/00031305.2017.1375989

Kako izgleda dobro organizirana tablica?

- Jedna vrijednost u jednoj ćeliji.
- Jedno zaglavlje s jasnim nazivima varijabli.
- Dosljedan zapis datuma, primjerice 2026-07-15.
- Jasno navedene mjerne jedinice.
- Dosljedno i dokumentirano označavanje nedostajućih vrijednosti.
- Bez oslanjanja na boju ćelije kao jedini nositelj informacije.
- Izvorni podaci odvojeni od izračuna i obrađenih verzija

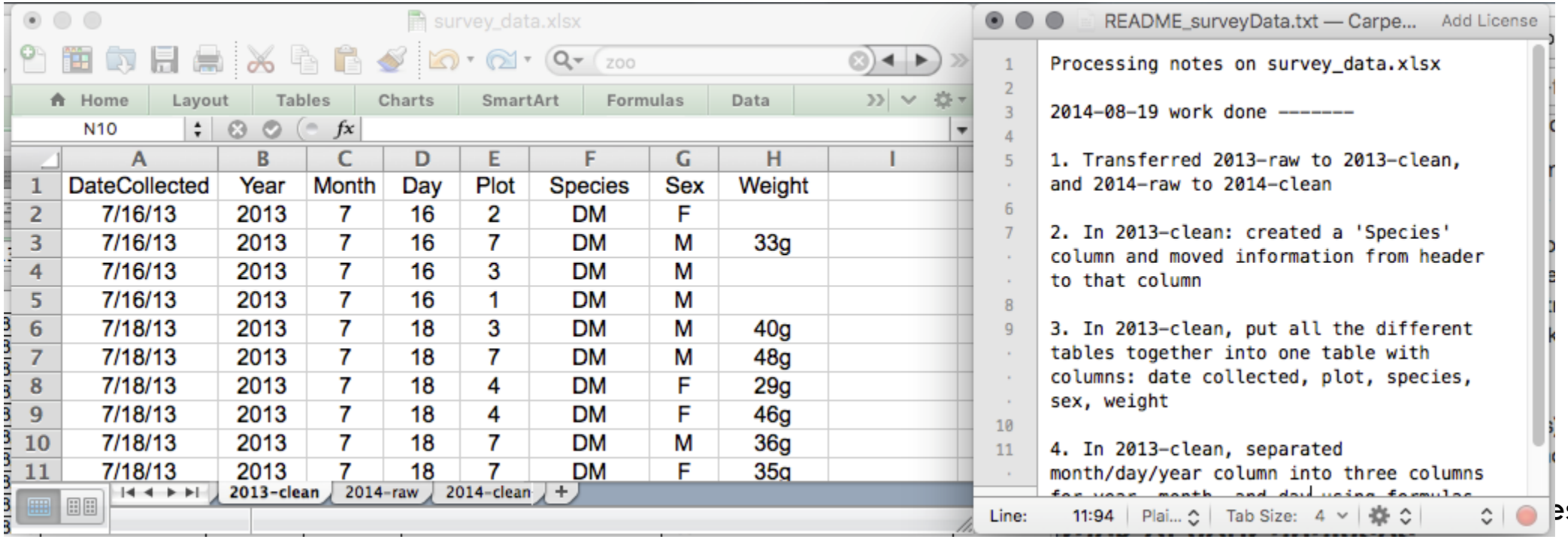
	A	B	C	D	E	F
1	mouse_id	sex	week	date	glucose	weight
2	3005	M	4	3/30/2007	19.3	635
3	3005	M	6	4/11/2007	31	460.7
4	3005	M	8	4/27/2007	39.6	530.2
5	3017	M	4	10/6/2006	25.9	202.4
6	3017	M	6	10/19/2006	45.1	384.7
7	3017	M	8	11/3/2006	57.2	458.7
8	3434	F	4	11/22/2006	26.6	238.9
9	3434	F	6	12/6/2006	45.9	378
10	3434	F	8	12/22/2006	56.2	409.8
11	3449	M	4	1/5/2007	27.5	121
12	3449	M	6	1/19/2007	42.9	191.3
13	3449	M	8	2/2/2007	56.7	182.5
14	3499	F	4	1/5/2007	19.8	220.2
15	3499	F	6	1/19/2007	36.6	556.9
16	3499	F	8	2/2/2007	43.6	446

Značenje oznake ne smijemo pretpostavljati; mora biti jasno definirano u dokumentaciji. Dosljedno označavanje nedostajućih vrijednosti i dokumentiranje njihova značenja dio su preporučene organizacije ekoloških tablica.

ORGANIZACIJA KOJA OMOGUĆUJE RAZUMIJEVANJE

Načela:

jasno definirana opažanja i varijable; dosljedni identifikatori uzoraka; dosljedni zapisi datuma i nedostajućih vrijednosti; dokumentirane mjerne jedinice; očuvani izvorni podaci; odvojena i razumljiva dokumentacija.



	A	B	C	D	E	F	G	H	I
1	DateCollected	Year	Month	Day	Plot	Species	Sex	Weight	
2	7/16/13	2013	7	16	2	DM	F		
3	7/16/13	2013	7	16	7	DM	M	33g	
4	7/16/13	2013	7	16	3	DM	M		
5	7/16/13	2013	7	16	1	DM	M		
6	7/18/13	2013	7	18	3	DM	M	40g	
7	7/18/13	2013	7	18	7	DM	M	48g	
8	7/18/13	2013	7	18	4	DM	F	29g	
9	7/18/13	2013	7	18	4	DM	F	46g	
10	7/18/13	2013	7	18	7	DM	M	36g	
11	7/18/13	2013	7	18	7	DM	F	35g	

esson/

KVALITETA PODATAKA?

Provjera	Primjer pitanja
Potpunost	Nedostaju li datumi, oznake uzoraka ili jedinice?
Dosljednost	Znače li „A”, „lokalitet A” i „postaja A” isto?
Raspon vrijednosti	Postoji li vrijednost koju treba provjeriti prema izvornom zapisu?
Dupliciranje	Je li isti zapis unesen dvaput?
Podrijetlo	Znamo li kako je pojedina vrijednost dobivena?
Status kvalitete	Je li vrijednost izmjerena, procijenjena ili označena kao upitna?

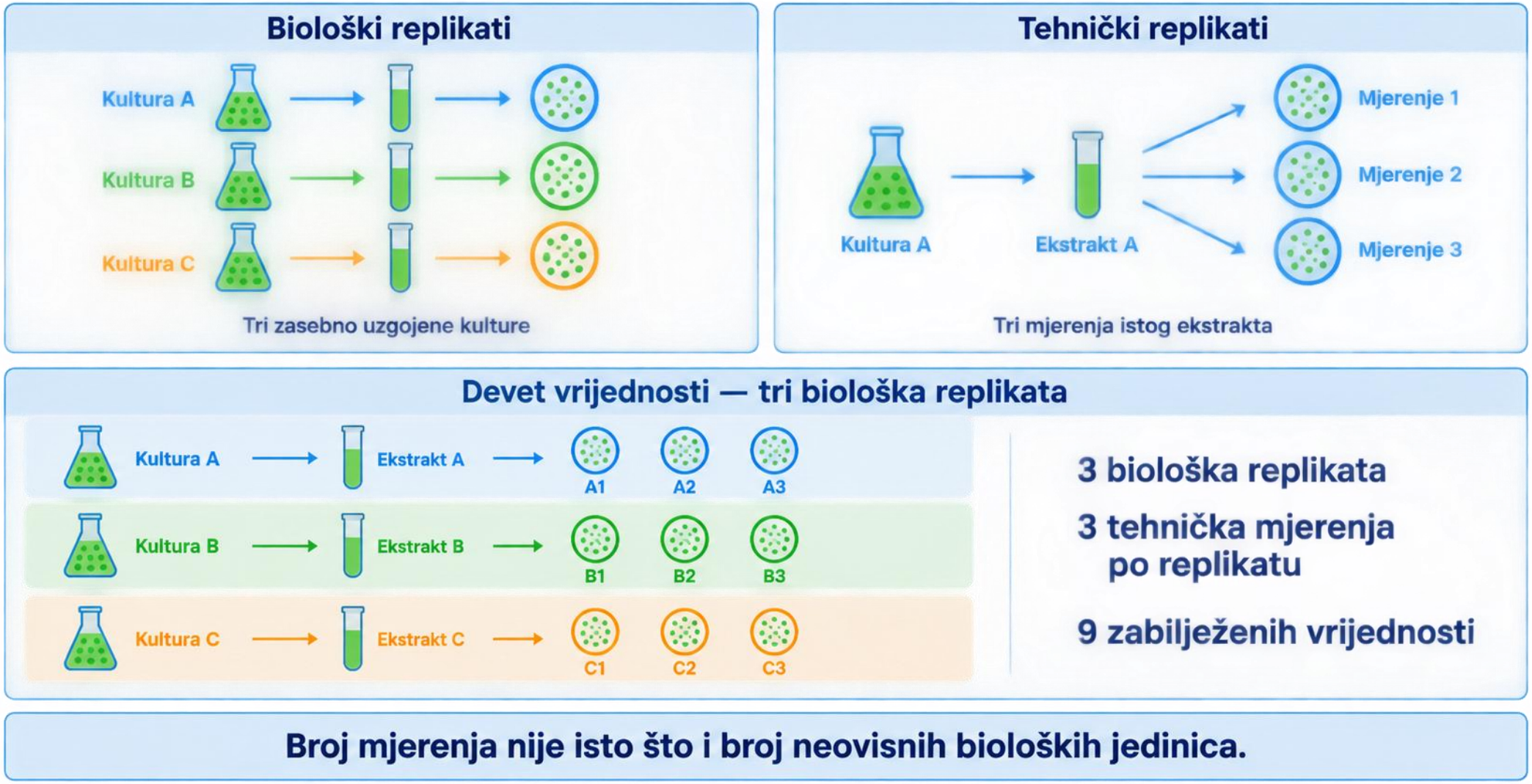
Sumnjivu vrijednost najprije provjeravamo i dokumentiramo; ne mijenjamo je samo zato što nam ne odgovara.

Nekoliko praktičnih pravila:

- sačuvati izvornu verziju podataka;
- jasno razlikovati izvorne i obrađene podatke;
- dokumentirati promjene i izvedene varijable;
- dosljedno zapisivati datume, jedinice i nedostajuće vrijednosti;
- izraditi kratak opis skupa podataka i rječnik varijabli

Možete li za šest mjeseci objasniti kako je od izvorne tablice nastao rezultat u vašem diplomskom radu?

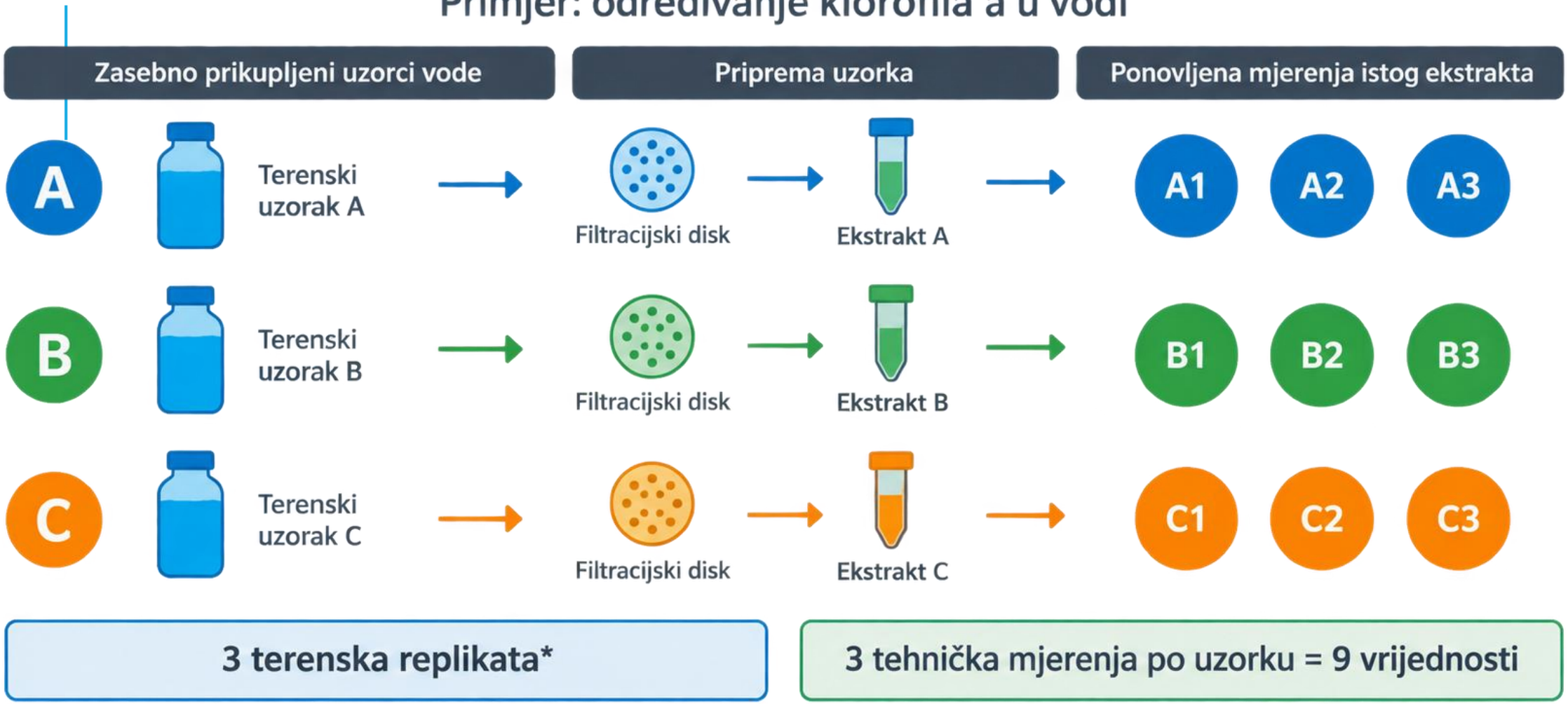
BIOLOŠKI I TEHNIČKI REPLIKATI



Ilustracija generirana AI alatom putem platforme Perplexity, 2026

Terenski uzorci i tehnička mjerenja

Primjer: određivanje klorofila a u vodi



*Neovisnost terenskih replikata ovisi o prostornom i vremenskom dizajnu uzorkovanja.
Tri uzorka s jednog lokaliteta nisu tri neovisna lokaliteta.

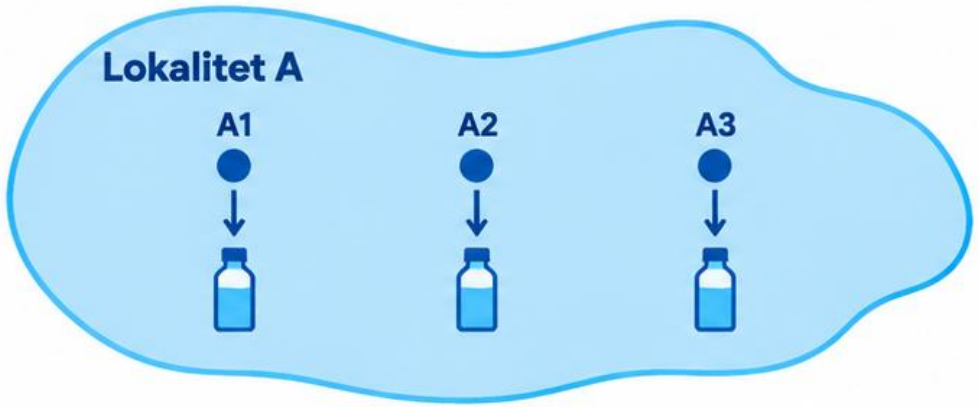
Ilustracija generirana AI alatom putem platforme Perplexity, 2026

https://davidzeleny.net/blog/2022/03/21/pseudoreplications-more-tricky-than-we-may-think/

Što želimo opisati ili usporediti?

Replikati se definiraju prema istraživačkom pitanju.

1. Opis jednog lokaliteta



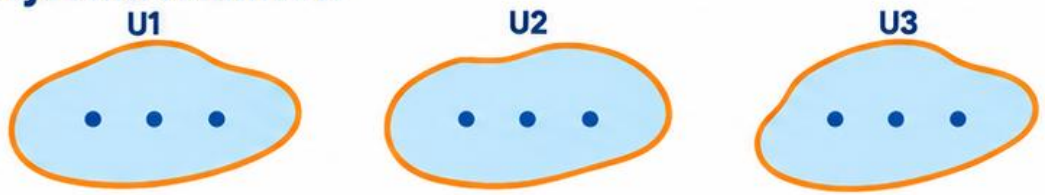
Tri terenska uzorka unutar jednog lokaliteta

Koliko se uzorci razlikuju unutar lokaliteta?

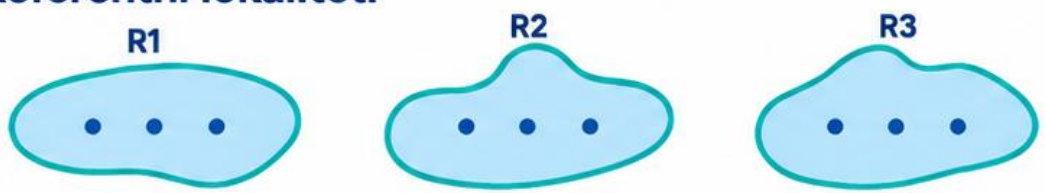
Ne predstavljaju tri lokaliteta.

2. Usporedba tipova lokaliteta

Utjecani lokaliteti



Referentni lokaliteti



Više lokaliteta po tipu; više uzoraka unutar lokaliteta.

Jedan lokalitet + deset uzoraka ≠ deset lokaliteta

Tip lokaliteta → **Lokalitet** → **Terenski uzorak** → **Poduzorak** → **Mjerenje**

Više mjerenja ne zamjenjuje više neovisnih lokaliteta.

Neovisnost uzoraka ovisi o prostornom i vremenskom dizajnu.

Ilustracija generirana AI alatom putem platforme Perplexity, 2026

ČITANJE PRIKAZA I INTERPRETACIJA

Redoslijed čitanja:

1. Što je na osima?
2. Koje su jedinice?
3. Što predstavlja točka, stupac ili linija?
4. Koje skupine uspoređujemo?
5. Je li prikazana pojedinačna vrijednost ili sažetak?
6. Koje informacije prikaz ne sadržava?

Grafička interpretacija sastavni je dio podatkovne pismenosti u biologiji.

ZADATAK 1

Uzorak	Datum	Lokacija	Temp.	Klorofil	Napomena
1	5/6/26	A	19,2	2,1	
2	06.05.2026.	lok. A	19.5	0	
3	2026-06-05	B	190	—	
3	2026-06-05	B	20,1	3,4	ponovljeno

Pronađite najmanje pet nejasnoća ili mogućih problema.
Za svaki navedite što biste provjerili prije analize.

ZADATAK 2

1. Navedite tri informacije bez kojih ne biste mogli pravilno interpretirati mjerenje koncentracije klorofila a.
2. Objasnite zašto nula i nedostajuća vrijednost ne moraju značiti isto.
3. Navedite jedan problem u tablici koji ne biste smjeli „ispraviti” bez provjere.

ZADATAK 3

Lokalitet A uzorkovan je u svibnju, a lokalitet B u kolovozu. Na svakom je lokalitetu prikupljen jedan uzorak vode. Koncentracija klorofila a određena je tri puta iz svakog uzorka.

Zaključak: Lokalitet B ima trajno višu koncentraciju klorofila a od lokaliteta A.

Lokalitet	Oznaka uzorka	Laboratorijsko mjerenje	Klorofil a
A	A01	1	1,2
A	A01	2	1,3
A	A01	3	1,2
B	B01	1	3,4
B	B01	2	3,5
B	B01	3	3,4

Što predstavlja svaki redak, a što oznaka uzorka?
Koje informacije nedostaju?
Koji dio zaključka treba posebno preispitati?
Kako biste preoblikovali zaključak da bude oprezniji?
Što biste promijenili u planu budućeg istraživanja?

Podaci trebaju biti upotrebljivi i nakon završetka istraživanja

FAIR načela:

Findable - podaci se mogu pronaći.

Accessible - podacima se može pristupiti pod jasno utvrđenim uvjetima.

Interoperable - podaci se mogu povezivati i koristiti s drugim podacima.

Reusable - podaci su dovoljno opisani i uređeni za ponovnu uporabu.

> Podatkovno pismen istraživač ne pita samo „Kako analizirati podatke?”, nego i „Što oni predstavljaju, kako su nastali i dokle sežu zaključci koje iz njih izvodimo?”