ALTERNATOR
Misliti znanost.

Veliko podatkovje in privid totalne zgodovine

6. 5. 2021

Number: 27/2021

Author:

- Dimitrij Mlekuž Vrhovnik



Foto: Arne Hodalič, računalniška obdelava: Katja Bidovec

Angleški arheolog O. G. S. Crawford je v začetku dvajsetega stoletja zgodovino primerjal s preprogo (https://www.cambridge.org/core/journals/antiquity/article/abs/historical-cycles/EBC0D0B538E2796AE7731CC1AAA1BAF1) , katere vzorce lahko prepoznamo le stoje, od zgoraj; za mačko, ki se sprehodi preko nje, preprogo sestavljajo le naključne barvne lise. Preproge zgodovine ne vidimo v celoti, zato ker nimamo dovolj empiričnih podatkov in pravega gledišča, od koder bi zaobjeli njeno totalnost. Arheologija je disciplina, ki preučuje človeško zgodovino v dolgoročni perspektivi skozi materialne ostanke; zanima jo preproga, ki smo jo stkali v zadnjih dveh milijonih let. Arheologija ima potencialno največji arhiv virov; vsi materialni ostanki globalne človeške preteklosti so potencialen vir.

Iz česa je preteklost?

Za arheologijo preteklost ne obstaja kot taka, temveč zgolj kot materialni ostanki v sedanjosti; ostanki, ki so se ohranili do danes, bodisi zaradi svoje trdoživosti, ugodnih pogojev bodisi zato, ker je nekdo zanje bolje skrbel. Ontološko preteklost obstaja kot sled; sled je materialni ostanek nekega dogodka ali procesa v preteklosti, ki je ostal do trenutka, ko ga zapišemo, odtisnjen v materialnem substratu. Večina sledi je nehotenih; pušča jih bivanje ljudi v materialnem svetu. Nekatere sledi so dobesedno sledi, stopinje (https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-0465-2_1891) , urezi kamnitega orodja na kosteh (https://royalsocietypublishing.org/doi/full/10.1098/rsfs.2016.0006) ; razmerja izotopov v kosteh (https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199569069.001.0001/oxfordhb-

9780199569069-e-8) pričajo o prehrani in mobilnosti, ohranjena DNK (https://link.springer.com/article/10.1007/s11245-019-09665-2) o populacijskih gibanjih. A zgodovina ni zgolj zgodovina človeštva, je zgodovina vpletanja ljudi v materialni svet. Preteklost je tako povsod; sledovi se vtisnejo v okolje; preteklost je zapisana v sedimentih (https://link.springer.com/chapter/10.1007/0-306-47654-1_12) , v razmerjih med pelodnimi zrni (https://www.sciencedirect.com/science/article/pii/B9780120031061500109) , v genskem zapisu (https://bmcbiol.biomedcentral.com/articles/10.1186/s12915-019-0724-7) vseh bitij na planetu; v obliki površja (https://www.jstor.org/stable/j.ctvh1dqdz) , v sestavi atmosfere (https://www.pnas.org/content/117/15/8250) … Nekatere sledi so več kot zgolj to; so stvari, izdelane z namenom, z načrtom; so sled obrazcev in idej, morda ubesedenih; namenjene aktivnemu poseganju v svet, oblikovanju vzorcev interakcij z okoljem ali med ljudmi: orodja, orožja, stavbe, naselbine, zemljiška razdelitev. Nekatere od sledi so simboli, stojijo namesto zapletenih družbenih konvencij. Tudi na te sledi se nalagajo sledi dejanske rabe, ki je pogosto drugačna kot zamišljena.

Arheološko rahločutnost zaznamuje zanimanje za vsakdanje. Vsaka sled je enako zanimiva. Pri brskanju skozi ruševine in smeti preteklosti je vse, dobesedno vse, potencialno zanimivo; vse lahko nosi informacijo o preteklosti. A množica potencialnih sledi še ni zgodovina. Sledi moramo najprej prepoznati, zapisati. Sledi niso še podatki; podatki postanejo skozi delo arheologov; stikanje po usedlinah preteklosti, prepoznavanje in zapisovanje pomenljivih sledi, urejanje zapisov. Če gledamo arheologe pri delu – na najdišču, v laboratoriju ali v pisarni – večino časa posvečajo zapisovanju, urejanju zapisov; sestavljanju tega, kar imenujemo arheološki zapis (https://www.cambridge.org/si/academic/subjects/archaeology/archaeological-science/understanding-archaeological-record) . Arheološki zapis je skupek vsega tistega, kar arheologi naredijo skozi delo; je eklektičen skupek materialnih ostankov (škatle najdb, ruševin, ki ostanejo na najdiščih; ampule z vzorci), zapisov, načrtov, fotografij, seznamov, meritev, podatkovnih zbirk … A ker večina preteklih dogodkov ne pusti sledov; ker je preteklosti v sedanjosti čedalje manj, ker sledovi preteklosti izginjajo, se izgubljajo, bledijo; ker jih kasnejši sledovi prekrijejo, izbrišejo; je arheološki zapis že v osnovi luknjičav, pomanjkljiv, razdrobljen. Arheološki zapis definirajo ravno ti negativi, luknje, praznina. To je prostor delovanja arheologije; če uporabim metaforo preproge, je delo arheologije prepoznavanje vzorcev kljub luknjam in vrzelim v preprogi, s pomočjo redkih ohranjenih in skrbno zapisanih niti.

Arheologijo zadnjih dveh desetletij zaznamuje prava eksplozija novih možnosti prepoznavanja sledi: DNK, razmerja izotopov v kosteh, ostanki maščob v črepinjah (https://onlinelibrary.wiley.com/doi/book/10.1002/9781444392449) , sledovi škroba v zobnih oblogah (https://www.nature.com/news/neanderthal-tooth-plaque-hints-at-meals-and-kisses-1.21593) … odpirajo popolnoma nove vidike preteklosti, za katere se je zdelo, da so za vedno izgubljeni. Arheološko delo, delo prepoznavanja in zapisovanja sledov, je drago in počasno. Podatki so razpršeni in raztreseni po različnih podatkovnih bazah, objavah, poročilih. Arheološki zapis ni le luknjičav in poln praznin, je tudi razpršen in raztreščen; po depojih muzejev, po arhivih inštitutov; po objavah, člankih, knjigah, po težko dostopnih poročilih, t. i. sivi literaturi (https://intarch.ac.uk/journal/issue40/6/toc.html) . A ker je vse, kar danes je, nastalo v preteklosti, je preteklost povsod; vsak podatkovni niz vsebuje tudi sledove preteklosti. Podatki daljinskega zaznavanja (https://www.springer.com/gp/book/9780387444536) , sistematičnega dokumentiranja Zemlje iz vesolja prinašajo ogromno podatkov o preteklosti, ki se kažejo kot vzorci rastja ali grbine in poglobitve v površju. Podatki, ki nastajajo pri gradbenih posegih, spremljanju okolja, pri geoloških preiskavah in preciznem kmetijstvu (https://link.springer.com/article/10.1007/s12520-017-0564-8) , fotografije (https://mw18.mwconf.org/paper/digital-tools-and-how-we-use-them-the-destruction-and-reconstruction-of-tangible-cultural-heritage/) , ki jih turisti naredijo na arheoloških najdiščih in naložijo na Instagram; vse to skriva tudi arheološke sledi. Tudi pridobivanje podatkov se ceni. Kmalu bodo po poljih križarila krdela majhnih avtonomnih robotkov (https://dl.acm.org/doi/10.5555/1659997.1660026) in štela in prepoznavala črepinje in merila fizikalne lastnosti tal in opozarjala na sledove preteklosti. Še dlje v prihodnosti bodo morda roji nanorobotov rili po prsti kot bakterije in za vsak kubični milimeter Zemljinega površja zabeležili kemično sestavo, fizikalne lastnosti in prisotnost DNK.

Izziv za arheologijo ni le prepoznavanje, kopičenje in tolmačenje sledov, temveč čedalje bolj tudi združevanje, zlivanje, kombiniranje vseh teh podatkov v velik arheološki zapis.

Veliko podatkovje

Arheologija se tako kot druge discipline sooča z izzivi velikega podatkovja (angl. *big data*). Veliko podatkovje so ogromni, nepredstavljivo veliki podatkovni nizi, ki se kopičijo z neznansko hitrostjo; prihajajo iz množice različnih virov in nastopajo v mnogih oblikah. Veliko podatkovje je pojav poznega kapitalizma; ko sledi človeštva ostajajo kot digitalno podatkovje, ki jih kapitalizem potrebuje za nov cikel rasti, ki temelji na nadzoru našega vedenja (https://www.publicaffairsbooks.com/titles/shoshana-zuboff/the-age-of-surveillance-capitalism/9781610395694/) . Zdi se, da je cilj velikega podatkovja izdelati zemljevid resničnosti v merilu 1 : 1, kot v znameniti zgodbi Jorga Luisa Borgesa; zemljevid, ki ga kapitalizem potrebuje (kot vsi kolonizatorji potrebujejo zemljevide), da z njim bolje nadzoruje resničnost. Veliko podatkovje je tako nova entiteta, s katero se soočamo kot človeštvo. Ta zemljevid v razmerju 1 : 1 spreminja realnost samo. Kaj pomeni za znanost, humanistiko, zgodovinopisje, arheologijo? Kakšna bo preteklost, napisana iz velikega podtakovja? Velikega podatkovja ne moremo razumeti le kot količine; kaj še majhno in kaj že veliko? Pravi epistemološki izziv velikega podatkovja je nova kvaliteta, ki vznikne iz te ogromne kvantitete. Izziv velikega podatkovja ni razvoj boljših in učinkovitejših orodij in metod za obvladovanje podatkov, za krčenje velikega podatkovja na obvladljivo, »majhno« velikost, temveč v razumevanju, kaj nova kvaliteta, ki pride z velikostjo, prinaša. Z velikim podatkovjem se moramo soočiti drugače. Če je za »malo podatkovje«, podatke, kot smo jih navajeni, značilna čistost; torej velika

natančnost, malo napak in visoka kvaliteta, saj so podatki zbrani za določen namen, na smiseln in učinkovit način; nad njimi imamo pregled in nadzor, je potemtken veliko podatkovje umazano: motno, polno redundanc, šumov in napak. Če je bil prej vse signal, informacija, ki jo potrebujemo za neki namen, je pri velikem podatkovju skoraj vse šum; šum, ki zakriva potencialno pomenljive informacije. Ta radikalni zdrs v razmerju med signalom in šumom nas postavlja v popolnoma asimetričen položaj do podatkovja. »Veliko« v velikem podatkovju je veliko v odnosu do nas, ljudi. Veliko podatkovje nas, ljudi, pomanjša. Ljudje s svojimi ubornimi telesi in omejenimi kognitivnimi sposobnostmi enostavno nismo opremljeni, da bi lahko zaobjeli entiteto, kot je veliko podatkovje, v vsem njenem obsegu, se z njo pomenljivo soočali in smiselno ovrednotili pomen teh interakcij. Veliko podatkovje je sublimno; ogromen, brezobličen in brezmejen eksces neskončno drobnih nepomembnih detajlov.

A če sprejmemo dejstvo, da je cunami velikega podatkovja pač ogromen, umazan in moten; če se upremo želji, da bi ga očistili in zreducirali na obvladljiv curek, lahko v zameno dobimo množico novih uvidov, novih pogledov in presenetljivih spoznanj. Veliko podatkovje nam omogoča uvide v popolnoma nova merila opazovanja, tako prostorska kot časovna. Odpira nam poglede, ki hkrati vključujejo celoto in mikrodetajle; odstira nam trende, vzdolžne ravni, vzorce odklonov od normalnega, anomalije. Omogoča nam uvide v razlike med merili, v razmerja med detajli in celoto, omogoča nam opazovanje vzorcev, ki nastajajo v interakcijah med različnimi merili opazovanja. Omogoča uvide v kompleksne procese in vznikanje oblik. Soočanje z motnim velikim podatkovjem zahteva drugačne načine dela s podatki. Pri velikem podatkovju ne obdelujemo posamičnih podatkov; pozorni smo na pojave, ki vzniknejo iz podatkovja: korelacije, trende, variabilnosti, anomalije, izjeme, vzorce. Zahteva nove načine analize podatkov: hierarhične reprezentacije, globoko učenje, različne vizualizacije … Veliko podatkovje zahteva, da velik del kognitivnega dela s podatki delegiramo ne-ljudem, različnim algoritmom, ki opravljajo »umazano« delo ponavljajočega se mukotrpnega prebijanja skozi podatke, preoblikovanja, filtriranja, računanja trendov, identifikacije anomalij in ekstremnih vrednosti ter prepoznavanja vzorcev. Ljudem ostane postavljanje pravih vprašanj, ukvarjanje s tem, kar prebavijo algoritmi, prepoznavanje in interpretacija zanimivih vzorcev in ukvarjanje z izjemami, odstopanji in skrajnimi vrednostmi. Hiter razvoj umetne inteligence, globokega učenja, konvolucijskih nevronskih mrež … je pripeljal do točke, ko algoritmi specializirane naloge prepoznavanja kompleksnih vzorcev (https://www.pnas.org/content/117/31/18240) , ki zahtevajo spretnosti in izkušnje, opravljajo bolje kot ljudje. Obeta se nam, da algoritmi kmalu ne bodo zgolj pomočniki v procesu produkcije znanja, temveč bodo aktivno ustvarjali znanje, morda hitreje in bolje kot ljudje. Veliko podatkovje lahko popolnoma spremeni način produkcije znanja o preteklosti in napiše zgodovino na novo.

Kakšno zgodovino prinaša veliko podatkovje?

A pričakovati, da se bo iz umazanega velikega podatkovja preteklosti kar sama napisala za nas, ljudi, pomenljiva zgodovina, je nerealno in nevarno. Nekritičen optimizem, da bodo ob dovolj številnih podatkih ti govorili sami zase; da je zemljevid preteklosti v merilu 1 : 1 že enak zgodovini, je zelo naiven. Najprej zato, ker podatki niso nikoli »surovi«, nevtralni, inertni; podatki so vedno »kuhani«, torej zbrani in pripravljeni skozi prakse zbiranja. In ker je arheološki zapis rezultat arheološkega dela, je poleg vira o preteklosti tudi vir o praksah, ideologijah, naracijah, teorijah in predpostavkah arheološke discipline. Poleg podatkov o preteklosti je veliko podatkovje hkrati tudi kolektivno nezavedno arheologije. Pri velikem podatkovju preteklosti še posebej arheologiji niti zemljevid v merilu 1 : 1 ne bo dovolj, saj je arheološki zapis v izhodišču luknjičav, pomanjkljiv in fragmentiran; teh vrzeli tudi veliko podatkovje ne bo zapolnilo. Preteklost izginja, to je v naravi preteklosti; arheologija okoli ostankov preteklosti tke podobe preteklosti. Arheologija se ukvarja s sledovi individualnega delovanja, ki so sami zase banalni in nesmiselni, a sestavljajo pomenljive vzorce v prostoru in času. Arheologija se ukvarja tako s senzorično specifičnostjo vsake sledi kot njenega mesta in učinka v konfiguraciji, ki se odvija v prostoru in času. Ti vzorci in pravilnosti vznikajo v prostoru in času, živijo in izginejo, se morda preoblikujejo v nekaj novega. To dinamiko preteklosti je arheologija pospravila v razumljive, vnaprej pripravljene antropocentrične obrazce, predalčke in kategorije, kot so kulture, kulturne skupine, tehnokompleksi, civilizacije. Te kategorije so razumljive, obvladljive in urejene v naracije, ki največkrat oblikujejo ideje o izvoru, o človeški izjemnosti, o razvoju in napredku ter o dihotomiji med naravo in kulturo. Algoritmi v velikem podatkovju iščejo predvsem vzorce; strojno učenje je prej način zaznave, orodje za iskanje skritih in neintuitivnih povezav, korelacij, kot orodje za odkrivanje vzročnih povezav, ki bi jih lahko generalizirali v zakone. Ko se podatkovje kopiči, gosti, ko se v njem izostrujejo detajli, začno v njem vznikati vzorci, pravilnosti, obrazci, strukture, neintuitivne in presenetljive strukture, ki zastavljajo nova vprašanja in nove smeri raziskovanja.

Veliko podatkovje preteklosti je predvsem priložnost za odkrivanje novih in presenetljivih vzorcev preteklosti. Lahko le slutimo, da ti vzorci v prostoru in času ne bodo majhne, urejene, ljudem razumljive stvari, temveč zapleteni skupki, mešanice ljudi in ne-ljudi, kjer ljudje igramo podrejeno vlogo. Pričakujemo lahko nepredstavljivo sublimne entitete, kot so recimo hiperobjekti (https://www.upress.umn.edu/book-division/books/hyperobjects) , stvari tako ogromne, tako razpršene v prostoru in času, da jih preprosto ne moremo več ločiti od njihovih učinkov. Veliko podatkovje pomeni konec antropocentrične preteklosti, enostavnih naracij o razvoju in napredku. Veliko podatkovje prinaša drugačno arheologijo: arheologija, ki bo sicer prepozno, a vendar, pomagala razumeti sedanjost; pojave kot je antropocen (https://www.tandfonline.com/doi/full/10.1179/2042458215Y.0000000022) , obdobje ko je človeštvo postaja geološki dejavnik. Arheologija se bo še vedno ukvarjala z iskanjem in prepoznavanjem vzorcev; a veliko podatkovje prinaša tudi radikalno drugačno arheološko prakso; nova raziskovalna vprašanja; nove velike naracije, novo refleksijo o produkciji znanja v arheologiji. Veliko podatkovje je priložnost za bolj relevantno arheologijo, ki bo pomagala razumeti, kako smo pristali tu, kjer smo. Zato veliko podatkovje zahteva predvsem velika vprašanja. In velika vprašanja zahtevajo velike teorije. Vsa znanost postaja, ob soočanju z velikim podatkovjem, znanost velikega podatkovja.

https://www.alternator.science/en/long/veliko-podatkovje-in-privid-totalne-zgodovine/