

Otvoreni pristup istraživačkim podacima

Radionica "Open Access - otvoreni pristup znanstvenim informacijama"
Fakultet elektrotehnike i računarstva u Zagrebu, 24. listopada 2012.

Marijana Glavica, mglavica@ffzg.hr
Knjižnica Filozofskog fakulteta Sveučilišta u Zagrebu
<http://knjiznica.ffzg.unizg.hr>

Sadržaj

- što jesu a što nisu istraživački podaci?
- zašto je važno sigurno pohranjivanje i slobodna razmjena podataka?
- gdje mogu naći slobodno dostupne podatke za svoje istraživanje?
- gdje mogu pohraniti svoje podatke i učiniti ih dostupnima?
- što trebam znati kad pohranjujem i objavljujem svoje podatke ili kad koristim podatke koje su proizveli drugi istraživači?

Istraživački podaci

- originalni / primarni / sirovi
- prikupljeni ili kreirani za potrebe istraživanja
- korišteni za analizu na temelju koje se donose zaključci

... definirani kontekstom i svrhom

- isti materijal u jednoj situaciji jest u drugoj nije istraživački podatak




- prikupljaju se ili kreiraju za određenu svrhu
- mogu biti prikupljeni i korišteni za jednu svrhu, a naknadno se iskoristiti za drugu

Primjeri istraživačkih podataka

- mjerenja u eksperimentima
- kodirani odgovori u anketama
- audio zapisi, video, fotografije
- odgovori u testovima
- transkripti intervjua
- bilješke (laboratorijske, terenske, dnevnici)
- artefakti
- primjerci vrsta
- računalno generirani podaci
- modeli, algoritmi, skripte, log-datoteke, simulacijski softver, sheme
- metodologija i postupak, standardne procedure, protokoli

Što nisu istraživački podaci?

- znanstvene publikacije
- analize, opisi, statistike, grafovi, tablice, činjenice ili zaključci koji se pojavljuju u znanstvenim publikacijama
- podaci državnih tijela i službi (osim ako se koriste u svrhu istraživanja!)

 **Scientific Paper**

● **Comparison of radiotherapy alone with radiotherapy plus hyperthermia in locally advanced pelvic tumours: a prospective randomised, multicentre trial.**

van der Zee J, Gonzalez Gonzalez D, van Rhoon GC, van Dijk JD, van Putten WL, Hart AA. Subdivision of Hyperthermia, Academic Medical Centre, Amsterdam, The Netherlands

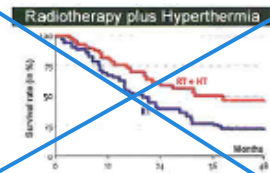
Background: Local-control rates after radiotherapy for locally advanced tumours of the bladder, cervix, and rectum are disappointing. We investigated the effect of adding hyperthermia to standard radiotherapy.

Methods: The study was a prospective, randomised, multicentre trial. 358 patients were enrolled from 1990 to 1996, in cancer centres in the Netherlands, who had bladder cancer stages T2, T3, or T4, N0, N1, cervical cancer stages IIB, IIIB, or IV, or rectal cancer stage M0-1 were assessed. Patients were randomly assigned radiotherapy (median total dose 65 Gy) alone (n=176) or radiotherapy plus hyperthermia (n=182). Our primary endpoints were complete response and duration of local control. We did the analysis by intention to treat.

Findings: Complete-response rates were 39% after radiotherapy and 55% after radiotherapy plus hyperthermia (p<0.001). The duration of local control was significantly longer with radiotherapy plus hyperthermia than with radiotherapy alone (p=0.04). Treatment effect did not differ significantly by tumour site, but the addition of hyperthermia seemed to be most important for cervical cancer, for which the complete-response rate with radiotherapy plus hyperthermia was 83% compared with 57% after radiotherapy alone (p=0.003). 3-year overall survival was 27% in the radiotherapy group and 51% in the radiotherapy plus hyperthermia group. For bladder cancer, an initial difference in local control disappeared during follow-up.

Interpretation: Hyperthermia in addition to standard radiotherapy may be especially useful in locally advanced cervical tumours. Studies of larger numbers of patients are needed for other pelvic tumour sites before practical recommendations can be made.

Published: Lancet 2000 Apr 1;355(9210):1119-25



Zašto je važno sigurno pohraniti podatke i učiniti ih dostupnima?

- digitalno doba - količina podataka sve je veća
- štednja resursa i izbjegavanje duplog posla
- mogućnost provjere rezultata istraživanja
- sekundarne analize - isti podaci drugi problemi
- povećava se šansa da rad bude korišten (i citiran)
- otežano falsificiranje podataka
- čuvanje podataka od gubitka i zaborava
- javnost može imati više povjerenja u znanstvena otkrića ako su podaci na kojima se temelje ta otkrića javno dostupni

Gdje su podaci?

- sve više sveučilišta i istraživačkih centara u svijetu gradi repozitorije podataka - pružaju trajan pristup podacima
- prisutna je velika heterogenost - zbog specifičnosti pojedinih područja
- otežano pronalaženje podataka

re3data.org

<http://www.re3data.org/>

- promovira kulturu razmjene podataka, teži ka povećanju vidljivosti istraživačkih podataka
- cilj je izraditi globalni registar repozitorija
- objava planirana za jesen 2012.
- prva faza:
 - konceptualizacija i konstrukcija registra
 - definicija kriterija za selekciju
 - formulacija metapodatkovnih schema za opis repozitorija - [Vocabulary for the Registration and Description of Research Data Repositories v1.0](#)

DataCite

<http://datacite.org/>

- internacionalni konzorcij koji se kao i re3data.org trudi da olakša pristup istraživačkim podacima
- razvija standarde za citiranje podataka (DOI)
- [Databib](#) - popis repozitorija istraživačkih podataka - radni dokument

Databib u facetama:

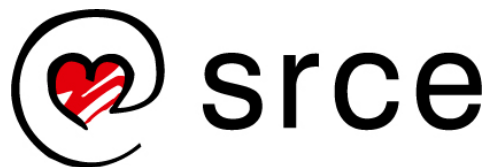
<http://webpac2.rot13.org/pub/exhibit/databib/databib.html>

Gdje mogu pohraniti podatke i učiniti ih dostupnima?

u Hrvatskoj

Javne kolekcije podataka

<http://www.srce.unizg.hr/proizvodi-i-usluge/podatkovne-usluge/jkp/>



u drugim zemljama

neki repozitoriji na Databib popisu prihvaćaju podatke autora iz različitih zemalja, neki su samo nacionalni

Koje formate mogu pohraniti?

- tekstualni (txt, pdf, rtf, html, xml)
- numerički (tablice ili objekti)
- multimedija (jpeg, tiff, mpeg)
- modeli (3D, statistički, makroekonomski, kauzalni)
- softver (programski jezici)
- specifični za disciplinu (FITS - astronomija, CIF - kemija, GRIB - meteorologija)
- specifični za instrument (model mikroskopa)

Tehnički aspekti

- poštovati standarde - ali koje? još su u razvoju
 - katalogizacija standarda u biomedicini [BioSharing](#) (Oxford)
- korisiti **slobodne formate**
 - osiguravamo mogućnost čitanja tih podataka u budućnosti (zbog razvoja tehnologije) !
 - olakšavamo razmjenu podataka

Kako prirediti podatke i gdje naći podršku?

- dokumentacija - opis metodologije, predmeta mjerenja, varijabli, ...
- metapodaci za opis skupa podataka - autor, naslov, godina, ...
- specifične prakse za različita područja
- edukacija o dobrim praksama priređivanja podataka za pohranjivanje i objavu

primjer: [MANTRA - Research Data Managemet Training](#)
(University of Edinburgh)

Licence

- u idealnom svijetu oslanjamo se na poštenje - citiranje originalnog autora - možemo li računati na to u stvarnom svijetu?
- ako se podacima ne pridruži licenca ne postoji ništa što kompetitivne i nepoštene može spriječiti da ukradu podatke

Licence za slobodne podatke

- za podatke nisu prikladne licence koje ograničavaju preradu ili omogućuju korištenje samo nekim korisnicima jer tako nije moguća potpuna sloboda pri integraciji podataka i njihovoj prenamijeni
- preporučaju se licence koje su u skladu s definicijom otvorenog znanja/otvorenih podataka
 - [Open Knowledge/Data Definition](#)
 - [Conformant Licenses](#)
 - [Panton Principles](#)

Ako vas ova tema stvarno zanima...

<https://www.zotero.org/mglavica/items/collectionKey/PXUGIFC6>