



Datenmanagement nach Rezept - Was Grossmutter's Backrezepte mit Forschungsdatenmanagement zu tun haben

Educational Material**Author(s):**

Hoffmann, André; Gloor, Daniel; [Petrus, Ana](#) ; Steurer, Florian; [Töwe, Matthias](#) 

Publication date:

2017-09-02

Permanent link:

<https://doi.org/10.3929/ethz-b-000297436>

Rights / license:

[Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International](#)

Datenmanagement nach Rezept

Was Grosis Backrezepte mit Forschungsdatenmanagement zu tun haben





Universität
Zürich ^{UZH}



ETH BIBLIOTHEK



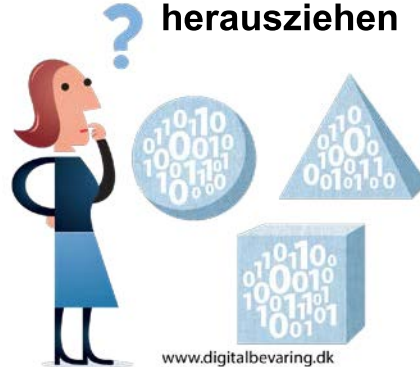
Was machen Forschende mit Daten?



Erheben, Messen,
Sammeln

www.digitalbevaring.dk

Information
erkennen und
herausziehen



www.digitalbevaring.dk



Schlüsse
ziehen

www.digitalbevaring.dk

Publiziertes
Wissen nutzen



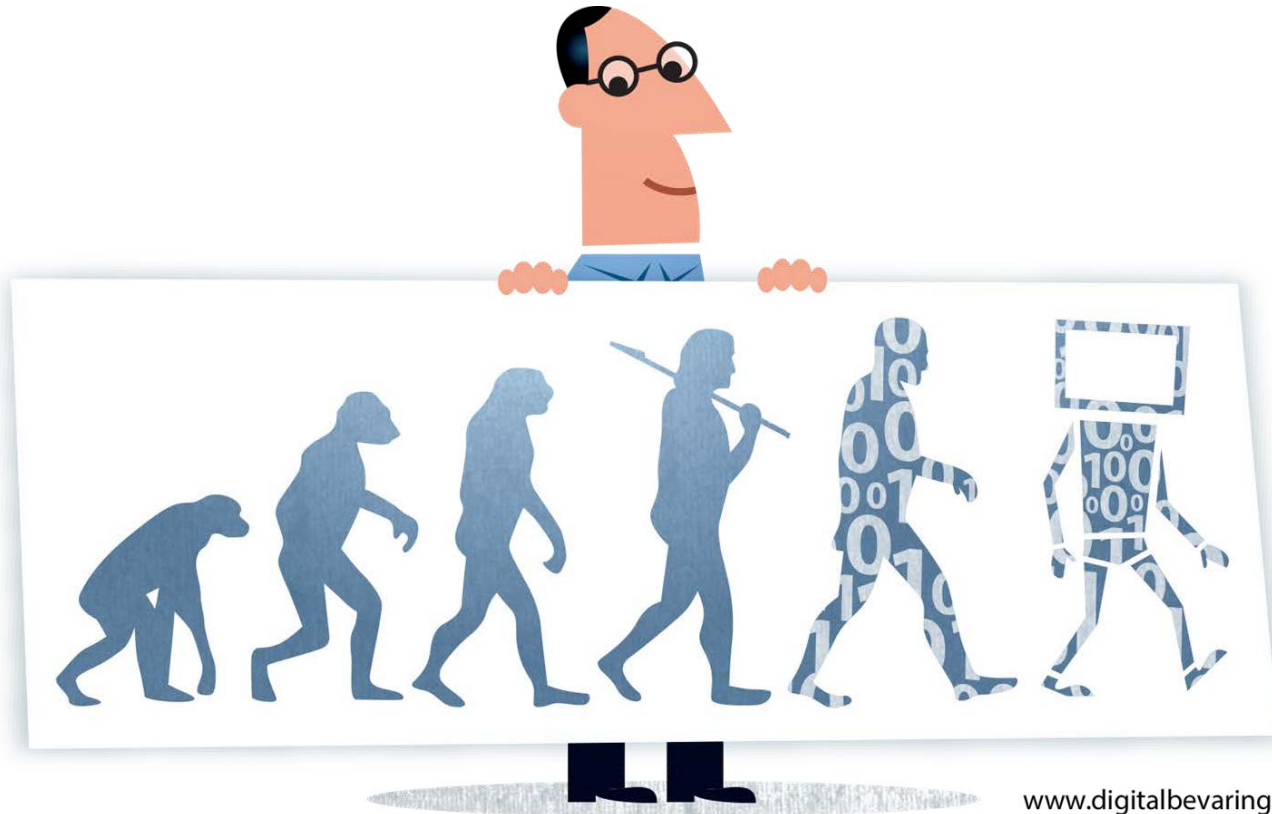
www.digitalbevaring.dk

Erkenntnisse
publizieren



www.digitalbevaring.dk

Aber Vorsicht - nicht alles ist digital...



...und jedes Fach arbeitet anders!



Was ist mit Forschungsdaten gemeint?

... in den Naturwissenschaften ...



Digital Curation Centre
Slide adapted from the PrePARE Project – [CC BY-SA](#)

Alle Daten, die Forschende bei ihrer Forschung nutzen oder erzeugen.

... und in Geisteswissenschaften

Your query "[(word="(P|p)arkier*")&(pos="(V.*ADJ.*))]" returned 666 matches in 514 different texts (in 575,285,200 words [1,699,115 texts]; frequency: 1.16 instances per million words) [0.553 seconds]

Show Page: 1 Line View Show in random order New query Go!

No	Filename	Solution 1 to 50	Page 1 / 14
1	2305773	gefeiert und war mit dieser in Streit geraten . Am Samstag gratis parkieren WEINFELDEN . Die Baustellen erschweren zurück	
2	2305773	im Zentrum . Es muss kein Geld eingeworfen werden , die Kunden parkieren länger als die erste Stunde gratis – das jeden Samst	
3	4550668	nun weitere Markierungen an . MARIO TESTA WEINFELDEN . Links und rechts parkierte Autos und Trottoirs ohne Niveaunterschied , dazu	
4	1149386	sucht Zeugen . Der Brand des bei der Kreuzung Längfeldw parkieren in der Nähe des Kreuzungspunktes am Samstag u	
5	2446916	wurde durch das Feuer vollständig zerstört . Die umliegende parkieren am Samstag u	

```

<div type="edition" xml:lang="la" xml:space="preserve"><head xml:lang="en">Text</head><ab
<lb n="1"/><persName type="attested"><name type="praenomen" nymRef="Caius"><expan><abbr
<persName type="attested"><name type="praenomen" nymRef="Caius"><expan><abbr
< lemma="filius"><expan><abbr>fil</abbr></ex
<name type="tribe" nymRef="Papiria"><w lemma
< lemma="flamen"><expan><abbr>flan</abbr></ex
<persName type="divine"><name nymRef="Liber">
< lemma="flamen"><expan><abbr>flan</abbr></ex
<persName type="attested"><name type="gentil
<lb n="2"/><w lemma="multa">multas</w> <w le
< lemma="exorno">exornavit</w> <w lemma="ac
< lemma="numero"><expan><abbr>n</abbr></ex>
<num value="12">XII</num>
<w lemma="exstruo">exstruo<lb n="3" type="inw
< lemma="statua">statuas</w> <w lemma="mar
< lemma="sestertius"><expan><abbr>am</abbr> <ex
<w lemma="mille"><expan><abbr>mil</abbr></ex>
<w lemma="nummus"><expan><abbr>num</abbr></ex>
<lb n="4"/><w lemma="aqua">aquae</w> <w lemma
<w lemma="promitto">promisit</w> <w lemma="e
<persName type="attested"><name type="cognom
<lb n="5"/><w lemma="ciuis">ciues</w> <w lemma
<w lemma="gladiator">gladiator</w> <w lemma
<w lemma="quinque"><expan><abbr>quin</abbr></ex>
<lb n="6"/><w lemma="splendidus">splendidiss
<w lemma="populus">populo</w> <w lemma="post
<w lemma="pono"><expan><abbr>ponend</abbr></ex>
<lb n="7"/><persName type="attested"><name t
<w lemma="honor">honore</w> <w lemma="conter
</ab></div>

```

Legend: ● parken, ■ parkieren

Word	Frequency
A.ille	~1
A.S.0	~2
D.Nominat	~4
D.Nominat	~5
D.S.0	~6
D.Nominat	~7
D.S.0	~8

Handwritten manuscript page with Latin text and a list of names: Mandel-Definitum, Prunillen-Definitum, Dypfel-Definitum, Eber-Definitum.

Was ist das Problem?



- In kürzerer Zeit werden immer mehr Daten erzeugt
- Der Druck, neue Ergebnisse zu präsentieren, ist hoch
- Forschen ist spannender als Daten zu ordnen...
- ...aber Datenmanagement lässt sich nicht nachholen!

Mit Forschungsdaten ist es wie mit Backrezepten... es kann viel schiefgehen

Kaffee**flecken**, Wasserschaden, Hauskatze **zerfetzt** das Rezept, es wird **verlegt**, die Handschrift ist **unleserlich**...



Harddisk wird **unleserlich**, der Laptop geht **verloren**, Daten wurden **nicht gesichert**, Files sind **veraltet** und lassen sich nicht mehr öffnen, der Doktorand nimmt Daten mit und hinterlässt keine Kopie, man vergisst wo man die Daten abgelegt hat...

Doch wieso sind Backrezepte und Forschungsdaten überhaupt so wichtig?

Man kann es **nachbacken**

Der “Output” ist **lecker**

Das gibt **Anerkennung** bei Freunden, wenn man es teilt

Und so das **kulturelle Erbe** der Grossmutter fortsetzt

Man kann die Forschung **replizieren** und **verifizieren**

Gewinnt dadurch **neue Erkenntnisse**

Und **Anerkennung** in der Forschungsgemeinschaft

Und **bringt** die **Menschheit** einen Schritt **vorwärts**

Was verraten “nackte” Daten...?

Daten ohne Zusammenhang sagen nicht viel bis gar nichts mehr aus!

- Woher kommen die Daten und wie wurden sie gewonnen (Bedingungen, Methoden usw.)?
- Wie muss / darf man sie interpretieren?
- (Wie) darf man sie verwenden?
- Bei wem kann man nachfragen?



**GARBAGE IN,
GARBAGE OUT!**

Wenn die Ausgangsdaten nicht klar sind, kann man daraus auch keine brauchbaren Schlüsse ziehen!

"MGB Grau Blau WP" (4.9.2018)
by Bidgee / CC BY-SA 3.0

Daten generieren und dokumentieren

Schlechte Backpraxis

Nehme Zutaten, backe bis goldgelb!

Apicius' antike "Käsepfanne":

7. Tyropatinam: Nimm Milch, **schätze die Menge** je nach Grösse der Auflaufform ab, **schmecke** die Milch mit Honig **wie für Milchbrei ab** und gib fünf Eier auf einen **Sextar**, wenn auf einen **Becher**, drei Eier dazu. Verrühre sie **so** in der Milch, **dass** du eine glatte Masse erhältst, passiere es in eine Tonschüssel und koch es auf **kleiner Flamme**. Wenn sie steif geworden ist, streue Pfeffer darüber und serviere.

Schlechte Forschungspraxis

Einfach einmal Chemikalien mischen!

Mische etwas Natrium mit Wasser!



Daten generieren und dokumentieren

Gute Backpraxis

200 g Mehl, 200 g Zucker, 2 Eier, 200 ml Milch

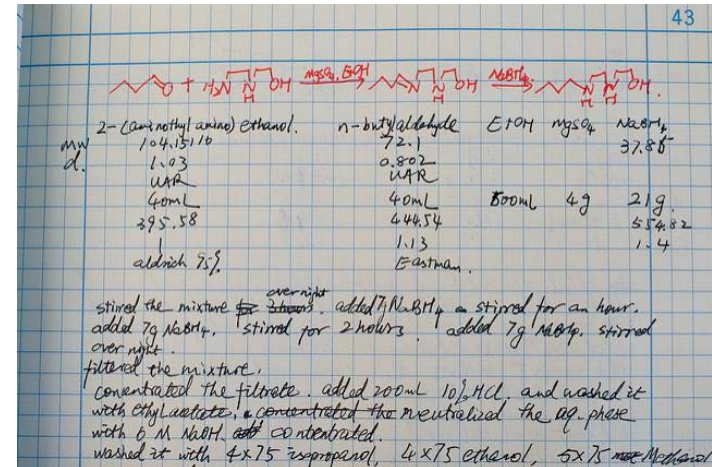
Mischen bis glatter Teig geformt

In mit Backpapier
belegter Form
bei 180 °C Umluft
20 min backen



Gute Forschungspraxis

genaue Mengenangaben, Laborprozedur,
Daten/Methode/Kontext dokumentieren



Daten ablegen und speichern

Gute Backpraxis

Im Rezeptheft leserlich aufgeschrieben

Grosi legt Rezept feinsäuberlich ab

auffindbar für Kinder und Enkel



Gute Forschungspraxis

“Sprechende” Filenamen vergeben

Datenträger systematisch beschriften

Daten standardisiert beschreiben und
verzeichnen

```
<mods:name usage="primary" type="personal">  
  <mods:namePart>Lavater, Johann Caspar</mods:namePart>  
  <mods:namePart type="date">1741-1801</mods:namePart>  
  <mods:role>  
    <mods:roleTerm type="code" authority="marcrelator">aut</mods:roleTerm>  
  </mods:role>  
</mods:name>
```

Daten finden und wiederverwenden

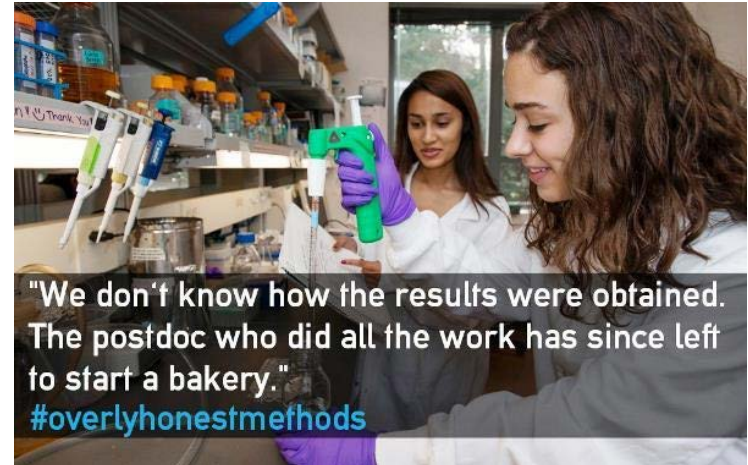
Schlechte Backpraxis

Rezept verschwindet bei der
Wohnungsräumung/Umzug



Schlechte Forschungspraxis

Doktorand nimmt Daten mit, Forschende
verschwinden nach Projektabschluss



Daten finden und wiederverwenden

Gute Backpraxis

Rezept ist in der Küche auffindbar

Rezept wird an nächste Generation weitergereicht



Gute Forschungspraxis

Daten verzeichnen und systematisch ablegen

Institution kann jederzeit auf Daten zugreifen



Warum kümmern sich Bibliotheken darum?

- Forschende bauen auf eigenem und fremdem Vorwissen auf:
„[...] *Didacus Stella*: „**Ein Zwerg, der auf den Schultern eines Giganten steht, wird weiter sehen können als der Gigant selbst**“
Robert Burton: *Anatomie der Melancholie* (1. Auflage 1621).
- Voraussetzung: Wissen ist auffindbar und zugänglich!



...weil sie ein Teil des Forschungsprozesses sind...

- Bibliotheken leisten dies für dokumentiertes Wissen seit Jahrtausenden...
- ...und immer stärker auch für die zugrundeliegenden Daten
- Junge Forschende verlassen eine Hochschule nach wenigen Jahren...
- ...während Institutionen wie Bibliotheken länger bestehen



...und es die Gesellschaft weiterbringt

- man muss das Rad nicht neu erfinden
- nicht replizierbare Daten werden bewahrt
- Bewahrung des wissenschaftlichen Erbes
- Citizen Science

- Beratung und Schulung zum Datenmanagement
- Research Collection zur Veröffentlichung von Forschungsdaten
- ETH Data Archive zur Langzeitarchivierung
- Registrierung von DOI (Digital Object Identifier) zum dauerhaften Auffinden
- Integration von ORCID (eindeutige Identifikation von Autorinnen)
- Open Access Publikation und Übernahme von Publikationsgebühren



- Data-Pilotprojekt mit HBZ und S3IT/Zentrale Informatik UZH
- Digitale Plattformen (e-manuscripta, e-rara in Zusammenarbeit mit ETH)
- Digitalisierungszentrum
- Bibliographische, standardisierte Datenerfassung
- Einträge in “Gemeinsamer Normdatei” (GND)



**Universität
Zürich** ^{UZH}

Hauptbibliothek

- Data-Pilotprojekt - Forschungsdatenmanagement und Infrastrukturen an der Universität Zürich
- DLCM - Aufbau von nationalen Standards und Dienstleistungen in Kooperation mit anderen Schweizer Hochschulen
- OpenAIRE 2020 - Unterstützung bei Open Access und Open Data für EU Projekte in der Schweiz
- ZORA - Zurich Open Repository and Archive

Take Home Message



Bildnachweise

Jørgen Stamp, <https://digitalbevaring.dk/>, CC BY 2.5 Denmark

Rezepte von Marija Popović, Ana Sesartić's Grossmutter

Bildcollage Forschungsdaten GW: [zu ergänzen]

Apic. de re coquinaria 7.13.7, deutsche Übersetzung: http://www.imperiumromanum.com/kultur/kulinarium/rezept_apicius07_13.htm

Chemistry, Paul, CC BY-NC-ND 2.0, <https://www.flickr.com/photos/howzey/4955405374/>

Kenyon_5.1.07_089, ydylg, CC BY-NC-ND 2.0, <https://www.flickr.com/photos/yutan/1269461280>

DATA, Janet McKnight, CC BY 2.0, <https://www.flickr.com/photos/janetmck/15422638442>

Piles of abandoned historical papers, Ralf Steinberger, CC BY 2.0, <https://www.flickr.com/photos/ralf-steinberger/32929137350>

Nell1, Lizbeth*King, CC BY-NC 2.0, <https://www.flickr.com/photos/10372146@N08/5729215336>

Research On Massive Amount of Climate Data Requires Fast Access, ibmphoto24, CC BY-NC-ND 2.0, https://www.flickr.com/photos/ibm_media/17408756412

Grandmother with grandson, Sasin Tipchai, <https://pixabay.com/photo-1822564/>, CC0

#overlyhonestmethods <https://i.imgur.com/QyJGd8O.jpg>

Recipes, George Oates, CC BY-NC-ND 2.0, <https://www.flickr.com/photos/george/5503973603>

Studierende, Bibliothek Erdwissenschaften, Ort: ETH Zürich, Zentrum, NO D 25.1, ETH Zürich / Alessandro Della Bella, 2013