Potentials and Limitations of Educational Datasets

Hendrik Drachsler
Open University of the Netherlands
Hendrik Drachsler

- **Assistant professor** at the Centre for Learning Sciences and Technologies (CELSTEC)
- Track record in **TEL projects** such as TENCompetence, SC4L, LTfLL, Handover, dataTEL.
- Main research focus:
  - **Personalization** of learning with information retrieval technologies, recommender systems and educational datasets
  - **Visualization** of educational data, data mash-up environments, supporting **context-awareness** by data mining
  - **Social and ethical implications** of data mining in education
- Leader of the **dataTEL Theme Team** of the STELLAR network of excellence (join the SIG on TELeurope.eu)
- Just recently: new **alterEGO project** granted by the Netherlands Laboratory for Lifelong Learning (on limitations of learning analytics in formal and informal learning)
dataTEL

Potentials and Limitations of Educational Datasets

24.07.2011 MUP/PLE lecture series, Knowledge Media Institute, Open University UK

Hendrik Drachsler
Centre for Learning Sciences and Technology
@ Open University of the Netherlands

#dataTEL
Goals of the lecture

1. Motivation or dataTEL
2. The dataTEL project
3. Potentials of dataTEL
4. Open issues of dataTEL
TEL RecSys Research

OpenScout

STELLAR

ReMashed
Recommendations for Mash-ups

Organic.Edunet

apoodle
learn @ work

VAA3R

TEN Competence

mavse
mining, analysis and visualization
based in social aspects of e-learning

funded by the Spanish Ministry of Science & Innovation,
Reference TIN2010-21715-C02-01
## Survey on TEL Recommender

### Table 12.3: Implemented TEL recommender systems reported in literature

<table>
<thead>
<tr>
<th>System</th>
<th>Status</th>
<th>Evaluator focus</th>
<th>Evaluation roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altered Vista [81, 82, 83, 104]</td>
<td>Full</td>
<td>Interface, Algorithm, System usage</td>
<td>Human users</td>
</tr>
<tr>
<td>RACOTH [2, 61]</td>
<td>Prototype</td>
<td>Algorithm</td>
<td>System designers</td>
</tr>
<tr>
<td>QMIA [26, 29]</td>
<td>Full</td>
<td>Algorithm</td>
<td></td>
</tr>
<tr>
<td>CYCLADES [4]</td>
<td>Full</td>
<td>Algorithm</td>
<td>System designers</td>
</tr>
<tr>
<td>Coltral [29, 30]</td>
<td>Prototype</td>
<td>System usage</td>
<td>Human users</td>
</tr>
<tr>
<td>Learning object sequencing [88]</td>
<td>Prototype</td>
<td>System usage</td>
<td>Human users</td>
</tr>
<tr>
<td>Evolving e-learning system [90, 91, 92, 93]</td>
<td>Full</td>
<td>Algorithm, System usage</td>
<td>Simulated users, Human users</td>
</tr>
<tr>
<td>EIS - Hybrid Personalised Recommender System [28]</td>
<td>Prototype</td>
<td>Algorithm, System usage</td>
<td>Human users</td>
</tr>
<tr>
<td>Multi-Attribute Recommendation Service [62]</td>
<td>Prototype</td>
<td>Algorithm</td>
<td>System designers</td>
</tr>
<tr>
<td>Learning Object Recommendation Model [95]</td>
<td>Design</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ReSearch [32]</td>
<td>Design</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation environment [72]</td>
<td>Full</td>
<td>Algorithm</td>
<td>Simulated users</td>
</tr>
<tr>
<td>BeMashed [26, 27]</td>
<td>Full</td>
<td>Algorithm, System usage</td>
<td>Human users</td>
</tr>
<tr>
<td>CourseRank [55, 56]</td>
<td>Full</td>
<td>System usage</td>
<td>Human users</td>
</tr>
<tr>
<td>CBR Recommender Interface [37]</td>
<td>Prototype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOSISE Recommendation Service [1]</td>
<td>Prototype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2M Recommending System [86]</td>
<td>Prototype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moodle Recommender System [44]</td>
<td>Prototype</td>
<td>Algorithm, System usage</td>
<td>Human users</td>
</tr>
<tr>
<td>LRLS [41]</td>
<td>Prototype</td>
<td>System usage, Learner performance</td>
<td>Human users</td>
</tr>
<tr>
<td>RTL recommender [49]</td>
<td>Prototype</td>
<td>System usage</td>
<td>System designers, Human users</td>
</tr>
</tbody>
</table>
### Table 12.3: Implemented TEL recommender systems reported in literature

<table>
<thead>
<tr>
<th>System</th>
<th>Status</th>
<th>Evaluator focus</th>
<th>Evaluation roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altered Vista</td>
<td>Full system</td>
<td>Interface, Algorithm, System usage</td>
<td>Human users</td>
</tr>
<tr>
<td>[81, 82, 82, 104]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RACOFI</td>
<td>Prototype</td>
<td>Algorithm</td>
<td>System designers</td>
</tr>
<tr>
<td>[2, 61]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QSAI</td>
<td>Full system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[78, 79]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYCLADES</td>
<td>Full system</td>
<td>Algorithm</td>
<td>System designers</td>
</tr>
<tr>
<td>[4]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CoFind</td>
<td>Prototype</td>
<td>System usage</td>
<td>Human users</td>
</tr>
<tr>
<td>[29, 30]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning object sequencing</td>
<td>Full system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[88]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIS – Hybrid Personal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendation System</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[28]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi-Attribute</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendation System</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[62]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning Object</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendation Ma</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[95]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RecSearch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[32]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation environment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[72]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ReMashed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[26, 27]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CourseRank</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[55, 56]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBR Recommender I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[33]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOSTLE Recommender Service</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[3]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEL Recommender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[90, 91, 92, 93]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Survey on TEL Recommender

Table 12.3: Implemented TEL recommender systems reported in literature

<table>
<thead>
<tr>
<th>System</th>
<th>Status</th>
<th>Evaluator focus</th>
<th>Evaluation roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altero Vista</td>
<td>Full system</td>
<td>Interface,</td>
<td>Human users</td>
</tr>
<tr>
<td>RACOTI[2, 61]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QAM</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observation:

Half of the systems (11/20) still at design or prototyping stage only 8 systems evaluated through trials with human users.

**Conclusion:**

Small-scale experiments with a few learners that rate some resources only adds little contributions to a knowledge base on recommender systems and personalization in TEL.

---

The TEL recommender research is a bit like this...
The TEL recommender research is a bit like this...

We need to design for each domain an appropriate recommender system that fits the goals, tasks, and particular constraints.
But...

“The performance results of different research efforts in TEL recommender systems are hardly comparable.”

(Manouselis et al., 2010)
But...

The TEL recommender experiments lack transparency. They need to be repeatable to test:

• Validity
• Verification
• Compare results
How others compare their recommenders
How others compare their recommenders

Although the TEL domain stores plenty of data everyday in e-learning environments (LMS, PLEs) there is a lack of shareable and publicly available datasets.
Goals of the lecture

1. Motivation or dataTEL
2. The dataTEL project
3. Potentials of dataTEL
4. Open issues of dataTEL
Who is dataTEL?

dataTEL is a Theme Team funded by the STELLAR network of excellence

Riina Vuorikari  Stephanie Lindstaedt  Katrien Verbert  Nikos Manouselis  Martin Wolpers  Hendrik Drachsler
Who is dataTEL?

dataTEL is a Theme Team funded by the STELLAR network of excellence

Riina Vuorikari  Stephanie Lindstaedt  Katrien Verbert  Nikos Manouselis  Martin Wolpers  Hendrik Drachsler

MAVSEL

Miguel Angel Sicillia  Joris Klerk

CEN PT Social Data
Who is dataTEL?

dataTEL is a Theme Team funded by the STELLAR network of excellence.

Riina Vuorikari
Stephanie Lindstaedt
Katrien Verbert
Nikos Manouselis
Martijn Wolpe

MAVSEL

Miguel Angel Sicilia
Joris Klerkx

CEN PT
Social Data
dataTEL::Objectives

Make the research on TEL RecSys more comparable by lowering the entrance barriers for other researchers and increase the quality.

The required benchmarks therefore are:

1. A collection of public available datasets ranging from formal to non-formal learning settings

2. An overview of the research results of certain RecSys technologies on different datasets

3. A common approach to evaluate RecSys in the domain of TEL
dataTEL::Objectives

1. Collecting publicly available datasets
2. Sharing policy to (re)use and share datasets
3. Define dataset standards (documentation, pre-processing)
4. Address privacy and legal protection rights
5. Create evaluation criteria for TEL recommender systems
6. Create a body of knowledge on personalization in TEL
dataTEL::Collection

RecSysTEL 2010
Workshop on Recommender Systems for Technology Enhanced Learning

Organised jointly by
4th ACM Conference on Recommender Systems (RecSys 2010)
5th European Conference on Technology Enhanced Learning (EC-TEL 2010)

Barcelona, Spain, 29-30 September 2010

Notifications to authors already sent.
Camera-ready versions have to be received by August 15, 2010

Submission for DataTEL challenge: August 31st, 2010 ***EXTENDED***

1st Workshop on Recommender Systems for Technology Enhanced Learning (RecSysTEL 2010)

In conjunction with

4th ACM Conference on Recommender Systems (RecSys 2010)
5th European Conference on Technology Enhanced Learning (EC-TEL 2010)

Barcelona, Spain, 29-30 September 2010

Keynote Speakers:

Joseph Konstan, GroupLens Research, University of Minnesota (USA)
# dataTEL::Collection

<table>
<thead>
<tr>
<th></th>
<th>Mendeley</th>
<th>APOSDELE</th>
<th>ReMashed</th>
<th>Organic.edunet</th>
<th>Mace</th>
<th>Melt</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Collection period</strong></td>
<td>1 year</td>
<td>3 months</td>
<td>2 years</td>
<td>9 months</td>
<td>3 years</td>
<td>6 months</td>
</tr>
<tr>
<td><strong>Users</strong></td>
<td>200,000</td>
<td>6</td>
<td>140</td>
<td>1,000</td>
<td>1,148</td>
<td>98</td>
</tr>
<tr>
<td><strong>Items</strong></td>
<td>1,857,912</td>
<td>163</td>
<td>96,000</td>
<td>11,000</td>
<td>12,000</td>
<td>1,923</td>
</tr>
<tr>
<td><strong>Activities</strong></td>
<td>4,848,725</td>
<td>1,500</td>
<td>23,264</td>
<td>920</td>
<td>461,982</td>
<td>16,353</td>
</tr>
<tr>
<td><strong>reads</strong></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td><strong>tags</strong></td>
<td>+</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>ratings</strong></td>
<td>(+)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>downloads</strong></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>search</strong></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td><strong>collaborations</strong></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>tasks/goals</strong></td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>sequence</strong></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
### dataTEL::Collection

<table>
<thead>
<tr>
<th></th>
<th>Mendeley</th>
<th>APOSDLE</th>
<th>ReMashed</th>
<th>Organic.edunet</th>
<th>Mace</th>
<th>Melt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collection period</td>
<td>1 year</td>
<td>3 months</td>
<td>2 years</td>
<td>9 months</td>
<td>3 years</td>
<td>6 months</td>
</tr>
<tr>
<td>Users</td>
<td>200,000</td>
<td>6</td>
<td>140</td>
<td>1,000</td>
<td>1.148</td>
<td>98</td>
</tr>
<tr>
<td>Items</td>
<td>1,857,912</td>
<td>163</td>
<td>96,000</td>
<td>11,000</td>
<td>12,000</td>
<td>1,923</td>
</tr>
<tr>
<td>Activities</td>
<td>4,848,725</td>
<td>1,500</td>
<td>23,264</td>
<td>920</td>
<td>461.982</td>
<td>16,353</td>
</tr>
<tr>
<td>reads</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>tags</td>
<td>+</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Collected data is very different with respect to amount of users and resources

Most of the data is very sparse

Privacy regulations harm data sharing

Mostly data from informal learning settings
Outcomes:
Tanimoto similarity + item-based CF was the most accurate.

Outcomes:
Tanimoto similarity + item-based CF was the most accurate.

Outcomes:
Implicit ratings like download rates, bookmarks can successfully used in TEL.


Fig. 3. F1 of user-based collaborative filtering with increasing number of neighbors
Goals of the lecture

1. Motivation or dataTEL
2. The dataTEL project
3. Potentials of dataTEL
4. Open Issues of dataTEL
Potentials of Open Data

Example by Tim Berners-Lee: The year open data went worldwide, TED talk FEB 2010
Potentials of Open Data

Example by Tim Berners-Lee: The year open data went worldwide, TED talk FEB 2010
Thousand years ago science was **empirical** (*Describing natural phenomena*)
Data = New Science Paradigm

• Thousand years ago science was **empirical** *(Describing natural phenomena)*

• Last few hundred years science: **theoretical** branch *(Using models, generalizations)*

\[ \Delta_n = \sum_{i=1}^{n} (-1)^{n+i} \cdot i^2 \]
Data = New Science Paradigm

• Thousand years ago science was **empirical** (*Describing natural phenomena*)

• Last few hundred years science: **theoretical** branch (*Using models, generalizations*)

• Last few decades: **computational** branch (*Simulating complex phenomena*)

\[ \Delta_n = \sum_{i=1}^{n} (-1)^{n+i} \cdot i^2 \]
Data = New Science Paradigm

• Thousand years ago science was **empirical** (*Describing natural phenomena*)

• Last few hundred years science: **theoretical** branch (*Using models, generalizations*)

• Last few decades: **computational** branch (Simulating complex phenomena)

• Nowadays: **data science** (*Unify theory, experiment, and simulation, data captured by instruments and processed by software, linked data*)

\[
\Delta_n = \sum_{i=1}^{n} (-1)^{n+i} \cdot i^2
\]
Promises of Open Data for TEL
Promises of Open Data for TEL

Unexploited potentials for TEL:

• The evaluation of learning theories and learning technology from the data side

• More transparent, mutually comparable, trusted and repeatable experiments that lead to evidence-driven knowledge

• Development of new educational data tools / products that combine different data sources in data mashups

• Gain new insights / new knowledge by combining so far unconnected resources / tools
Data Products

Explore flu trends around the world

We've found that certain search terms are good indicators of flu activity. Google Flu Trends uses aggregated Google search data to estimate flu activity. Learn more »
Data Products
Data Products

Educational Data Products

• Drop-out Analyzer
• Group Formation Recommender
• Question-Answering Tool
• Awareness Tools

Goals of the lecture

1. Motivation or dataTEL
2. The dataTEL project
3. Potentials of dataTEL
4. Open issues of dataTEL
dataTEL::Open issues

1. Privacy
2. Prepare datasets
3. Share datasets
4. Body of knowledge
Privacy

Listing all those empty homes out there

Also follow our twitter feed @pleaserobme.

Filter
Location
Twitter username

More Info
Home
Why
About

Made Possible By
Foursquare
Twitter
@boyvanamstel
Recent Empty Homes

@wgirarde left home and checked in less than a minute ago:

@TennesseeTimHi left home and checked in less than a minute ago:
Feeding my addiction ( @ Starbuck's) http://4sq.com/72m9vj
Privacy

OVERSHARING
Were the founders of PleaseRobMe.com actually allowed to take the data from the web and present it in that way?

<table>
<thead>
<tr>
<th>Recent Empty Homes</th>
</tr>
</thead>
<tbody>
<tr>
<td>@wgirarde left home and checked in less than a minute ago: Aplicativo para iPhone adiciona fotos e videos aos check-ins do Foursquare. &gt;&gt;&gt; <a href="http://bit.ly/dzIrN3">http://bit.ly/dzIrN3</a></td>
</tr>
</tbody>
</table>

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>@TennesseeTimHill left home and checked in less than a minute ago: Feeding my addiction (@ Starbucks's) <a href="http://4sq.com/72m9vj">http://4sq.com/72m9vj</a></td>
</tr>
</tbody>
</table>

Privacy
Privacy

OVERSHARING

Were the founders of PleaseRobMe.com actually allowed to take the data from the web and present it in that way?

Are we allowed to use data from social services and reuse it for research purposes?
Privacy
Privacy

1. Privacy as confidentiality
   The right to be let alone (Warren and Brandeis, 1890)
Privacy

1. Privacy as confidentiality
   The right to be let alone (Warren and Brandeis, 1890)

2. Privacy as control
   The right of the individual to decide what information about herself should be communicated to others and under which circumstances.
Privacy

1. Privacy as confidentiality
   The right to be let alone (Warren and Brandeis, 1890)

2. Privacy as control
   The right of the individual to decide what information about herself should be communicated to others and under which circumstances.

3. Privacy as practice
   The right to intervene in the flows of existing data and the re-negotiation of boundaries with respect to collected data.
Privacy solutions
Privacy solutions

1. Privacy as confidentiality
   Information services that minimizing, secure or anonymize the collected information
Privacy solutions

1. Privacy as confidentiality
   Information services that minimizing, secure or anonymize the collected information

2. Privacy as control
   Identity Management Systems (IDMS), with access control rules
Privacy solutions

1. Privacy as confidentiality
   Information services that minimizing, secure or anonymize the collected information

2. Privacy as control
   Identity Management Systems (IDMS), with access control rules

3. Privacy as practice
   Timestamp on data, data degradation technologies
Prepare datasets

Justin Marshall, Coded Ornament by rootoftwo
http://www.flickr.com/photos/rootoftwo/267285816
1. Create a dataset that realistically reflects the variables of the learning setting.
Prepare datasets

1. Create a dataset that realistically reflects the variables of the learning setting.

2. Use a sufficiently large set of user profiles

Justin Marshall, Coded Ornament by rootoftwo
http://www.flickr.com/photos/rootoftwo/267285816
Prepare datasets

1. Create a dataset that realistically reflects the variables of the learning setting.

2. Use a sufficiently large set of user profiles

3. Create datasets that are comparable to others

Justin Marshall, Coded Ornament by rootoftwo
http://www.flickr.com/photos/rootoftwo/267285816
Prepare datasets

For informal data sets:

1. Collect data
2. Process data
3. Document data
4. Share data

For formal data sets from LMS:

1. Data storing scripts
2. Anonymisation scripts
3. Document data
4. Share data
Prepare datasets

Info

About the Data Seal of Approval (DSA)

The Data Seal of Approval ensures that in the future, research data can still be processed in a high-quality and reliable manner, without this entailing new thresholds, regulations or high costs. The Data Seal of Approval and its quality guidelines may be of interest to research institutions, organizations that archive data and to users of that data. It can be granted to any repository that applies for it via the assessment procedure.

Anyone who archives his or her data would like to be able to find, recognise and use it in the future. With electronic data this cannot be taken for granted, after all hardware and software are changing all the time. Making data future-proof can be accomplished by ensuring that data sets and metadata meet certain requirements. In consultation with large data producers and managers, the Dutch data archive DANS documented what those requirements need to be in the Data Seal of Approval, which have been further developed since and handed over to the DSA Board in May 2009.

The quality guidelines of this seal of approval for data are intended to ensure that in the future, research data can still be processed in a high-quality and reliable manner, without this entailing new thresholds, regulations or high costs. The guidelines may be of interest to research institutions, organizations that archive data and to users of that data. You can download the document that contains the quality guidelines of the Data Seal of Approval below.

Contact: info@datasealofapproval.org
Share/cite datasets

Helping you to find, access, and reuse data

DataCite

What is DataCite?

We are a not-for-profit organisation formed in London on 1 December 2009. Our aim is to:

- establish easier access to scientific research data on the Internet
- increase acceptance of research data as legitimate, citable contributions to the scientific record
- support data archiving that will permit results to be verified and re-purposed for future study.

Our structure

Just as research is global, DataCite is global, with member institutions offering services and advice directly where they are needed by the researchers.

We consist of a Managing Agent, currently the German National Library of Science and Technology, with Members and Associate Members around the world.

Why cite data?

What is DataCite?

What do we do?

Email updates

Register for updates from 2013-05-01

Privacy by SafeSubscribers
Sharing policies
Sharing policies

Thank you for requesting Yahoo! Webscope data. Here are the steps to take to receive the data.
#1. Read the data sharing agreement below and if you agree, click "agree"
#2. You must also obtain your department chair approval to receive this data.
#3 Your department chair that you provided will receive an email and link to the data sharing agreement. They will need to "agree" to the data sharing agreement before we can provide the data. If you have any questions please contact us at research-data-requests@yahoo-inc.com

Agreement for Datasets:
R5. Yahoo! Delicious Popular URLs and Tags, version 1.0,
L9. Yahoo! Answers Question Types Sample of 1000, version 1.0

YAHOO! WEBSCOPE™ DATA LICENSE AGREEMENT
INTERNATIONAL

This data license agreement ("Agreement") is between you and Yahoo! Inc. ("Yahoo!") regarding your access to and use of the Yahoo! Webscope™ data that you wish to use and download from the Yahoo! website located at http://webscope.sandbox.yahoo.com ("Data"). Your access to and use of the Data is subject to the terms and conditions of this Agreement.

BY CLICKING THE "I AGREE" BUTTON, DOWNLOADING OR USING THE DATA, YOU AGREE THAT THAT YOU HAVE READ AND UNDERSTAND THE TERMS OF THIS AGREEMENT, AND THAT YOU AGREE TO BE BOUND AND TO ABIDE BY THIS AGREEMENT AND YAHOO'S TERMS OF USE. IF YOU DO NOT UNDERSTAND THIS AGREEMENT, DO NOT DOWNLOAD, ACCESS OR USE THE DATA.

1. SCOPE OF AGREEMENT
1.1 The Data is valuable and confidential information of Yahoo! You agree to use the Data only in accordance with this Agreement, and to hold the Data in strict confidence. You will not perform any analysis, reverse engineering or experimentation...
Thank you for requesting Yahoo! Webscope data. Here are the steps to take to receive the data.

1. Read the data sharing agreement below and if you agree, click "agree"
2. You must also obtain your department chair approval to receive this data.
3. Your department chair that you provided will receive an email and link to the data sharing agreement before we can provide the data. If you have any questions please contact us at research-data-requests@yahoo-inc.com.

Agreement for Datasets:
R5. Yahoo! Delicious Popular URLs and Tags, version 1.0,
L9. Yahoo! Answers Question Types Sample of 1000, version 1.0

YAHOO! WEBSCOPE™ DATA LICENSE AGREEMENT
INTERNATIONAL

This data license agreement ("Agreement") is between you and Yahoo! Inc. ("Yahoo!") regarding your access to and use of the Yahoo! Webscope™ data that you wish to use and download from the Yahoo! website located at http://webscope.sandbox.yahoo.com ("Data"). Your access to and use of the Data is subject to the terms and conditions of this Agreement.

BY CLICKING THE "I AGREE" BUTTON, DOWNLOADING OR USING THE DATA, YOU AGREE THAT THAT YOU HAVE READ AND UNDERSTAND THE TERMS OF THIS AGREEMENT, AND THAT YOU AGREE TO BE BOUND AND TO ABIDE BY THIS AGREEMENT AND YAHOO’S TERMS OF USE. IF YOU DO NOT UNDERSTAND THIS AGREEMENT, DO NOT DOWNLOAD, ACCESS OR USE THE DATA.

1. SCOPE OF AGREEMENT
1.1 The Data is valuable and confidential information of Yahoo! You agree to use the Data only in accordance with this Agreement, and to hold the Data in strict confidence. You will not perform any analysis, reverse engineering or processing.
Sharing policies

Next Steps:

- Yahoo! will review your request
- Your department chair will be contacted for their approval
- Upon approval of the data sharing agreement, Yahoo! will send an e-mail notification
- For those datasets that are not available via download we will ship those to the mailing address provided
- Data must be destroyed after 2 years

"agree" to the data sharing agreement before we can provide the data. If you have any questions please contact us at research-data-requests@yahoo-inc.com

Agreement for Datasets:
R5. Yahoo! Delicious Popular URLs and Tags, version 1.0,
L9. Yahoo! Answers Question Types Sample of 1000, version 1.0

YAHOO! WEBSCOPE™ DATA LICENSE AGREEMENT
INTERNATIONAL

This data license agreement ("Agreement") is between you and Yahoo! Inc. ("Yahoo!") regarding your access to and use of the Yahoo! Webscope™ data that you wish to use and download from the Yahoo! website located at http://webscope.sandbox.yahoo.com ("Data"). Your access to and use of the Data is subject to the terms and conditions of this Agreement.

BY CLICKING THE "I AGREE" BUTTON, DOWNLOADING OR USING THE DATA, YOU AGREE THAT THAT YOU HAVE READ AND UNDERSTAND THE TERMS OF THIS AGREEMENT, AND THAT YOU AGREE TO BE BOUND AND TO ABIDE BY THIS AGREEMENT AND YAHOO!’S TERMS OF USE. IF YOU DO NOT UNDERSTAND THIS AGREEMENT, DO NOT DOWNLOAD, ACCESS OR USE THE DATA.

1. SCOPE OF AGREEMENT
1.1 The Data is valuable and confidential information of Yahoo!. You agree to use the Data only in accordance with this Agreement, and to hold the Data in strict confidence. You will not perform any analysis, reverse engineering or decompilation of any kind on the Data.
Sharing policy guidelines

Do you want to publish some of the data that you are borrowing or to share it with someone other than your own team of researchers?
Then you probably need to secure the author’s consent.

Do you intend using a government database?
You do not need to seek consent for this.
Explanation: Unless provided otherwise, government databases are basically not subject to database right (Section 3.2.6). You do not require the consent of the rightholder in order to retrieve substantial portions of a government database in the framework of scientific/scholarly research (Section 3.2.5)

A brief guide on data licenses developed by SURF and the Centre for Intellectual Property Law (CIER), 2009 available at www.surffoundation.nl
Body of knowledge

Formal  Datasets  Informal

Data A  Data B  Data C

**Algorithms:**
- Algoritmen A
- Algoritmen B
- Algoritmen C

**Models:**
- Learner Model A
- Learner Model B

**Measured attributes:**
- Attribute A
- Attribute B
- Attribute C

**Algorithms:**
- Algoritmen D
- Algoritmen E

**Models:**
- Learner Model C
- Learner Model E

**Measured attributes:**
- Attribute A
- Attribute B
- Attribute C

**Algorithms:**
- Algoritmen B
- Algoritmen D

**Models:**
- Learner Model A
- Learner Model C

**Measured attributes:**
- Attribute A
- Attribute B
- Attribute C
Body of knowledge
Design experiments

This database lists resources about an approach to system development in which research findings are collated and distilled into best practice design - an engineering approach to TEL system development

To add a record, click on the plus symbol in the corner of the table. If you want to insert a link please go to our instruction page to see the syntax for the implementation of an external link.

If you would like to discuss a particular record in the comment section below the database, please first indicate which record you are discussing by citing the red coloured number on the left side of the table.

Search for resources using field below.

<table>
<thead>
<tr>
<th>Title/Author(s)/Date</th>
<th>Description</th>
<th>Type/URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>An impoverished machine: Challenges to human learning and instructional technology Taraban, R.(2008)</td>
<td>Behavior Research Methods paper on need to inform instructional technology based approaches to thermodynamics teaching from practices in the discipline</td>
<td>journal</td>
</tr>
<tr>
<td>Translating Constructivism into Instructional Design: Potential and Limitations Kozulin, V.</td>
<td>J. Educational Technology &amp; Society journal Discusses how 'moderate' versions of constructivist approaches</td>
<td>journal</td>
</tr>
</tbody>
</table>
Welcome to Linked Education

LinkedEducation.org is an open platform aimed at further promoting the use of Linked Data for educational purposes.

While sharing and reusing educational data across institutional and national boundaries is a general goal for both the public and the private education sector, the last decade has seen a large amount of research dedicated to Web-scale interoperability. The Linked Data approach – based on W3C standards such as RDF and a set of successful principles – has evolved as the general standard for sharing data on the Web and led to vast amounts of publicly available data sets, and a multiplicity of Linked Data-consuming applications which provide added-value services. However, there is only limited take-up of Linked Data principles in the educational field.
Body of knowledge
data TEL::SIG

http://www.teleurope.eu/pg/groups/9405/datatel/

Objectives:

- Representing dataTEL researchers to promote the release of open datasets from educational providers
- Fostering the standardizations of datasets to enable exchange and interoperability
- Contributing to policies on ethical guidelines (privacy and legal protection rights)
- Fostering a shared understanding of evaluation methods in TEL RecSys and Learning Analytics technologies.
Many thanks for your interests
Many thanks for your interests

Free the data

picture by Tom Raftery  http://www.flickr.com/photos/traftery/4773457853/sizes/l
Many thanks for your interests

This slide is available at:
http://www.slideshare.com/Drachsler

Email: hendrik.drachsler@ou.nl

Skype: celstec-hendrik.drachsler

Blogging at: http://www.drachsler.de

Twittering at: http://twitter.com/HDrachsler