Using Linked Data in Learning Analytics
LAK 2013 tutorial

Evaluation of Linked Data tools for Learning Analytics

Hendrik Drachsler (@hdrachsler, drachsler.de)
(CELTEC, Open Universiteit Nederland, NL)

Eelco Herder
(L3S Research Center, DE)

Mathieu d’Aquín (@mdaquín, mdaquin.net)
(Knowledge Media Institute, The Open University, UK)

Stefan Dietze
(L3S Research Center, DE)
What are the evaluation criteria of Robot Wars?

Criteria:
• Damage
• Aggression
• Control
• Applause
RecSysTEL Evaluation criteria

1. Accuracy
2. Coverage
3. Precision
4. Recall

1. Effectiveness of learning
2. Efficiency of learning
3. Drop out rate
4. Satisfaction

Combine approach by
Drachsl er et al. 2008

Kirkpatrick model by
Manouselis et al. 2010
Conclusions:

Half of the systems (11/20) still at design or prototyping stage only 9 systems evaluated through trials with human users.
The TEL recommender research is a bit like this...

We need to design for each domain an appropriate recommender system that fits the goals and tasks.
TEL recommender experiments lack transparency and standardization. They need to be repeatable to test:

- Validity
- Verification
- Compare results

http://www.flickr.com/photos/kaptainkobold/3203311346/
Data-driven Research and Learning Analytics

EATEL-Special Interest Group

Hendrik Drachsler (a), Katrien Verbert (b)

(a) CELSTEC, Open University of the Netherlands
(b) Dept. Computer Science, K.U.Leuven, Belgium
1st Workshop on Recommender Systems for Technology Enhanced Learning (RecSysTEL 2010)

Issues and Considerations regarding Sharable Data Sets for Recommender Systems in Technology Enhanced Learning

Hendrik Drachsler*, Toine Bogersb, Riina Vuorikari, Katrien Verbertc, Erik Duvald, Nikos Manouselis, Guenter Beham, Stephanie Lindstaedtf, Hermann Sternf, Martin...
Dataset-driven Research for Improving Recommender Systems for Learning

Katrien Verbert1, Hendrik Drachsler2, Nikos Manouselis3, Martin Wolpers4, Riina Vuorikari5 and Erik Duval1

1 Department of Computer Science, K.U.Leuven
Celestinlaan 200A, B-3001 Leuven, Belgium
{katrien.verbert, erik.duval}@cs.kuleuven.be
2 Open University of the Netherlands (OUNL)
P.O. Box 2900, 6400 DI Heerlen, The Netherlands
hendrik.drachsler@ou.nl
3 Greek Research and Technology Network (GRNET)
56 Massoglion Av., 115 27, Athens, Greece
nikos@grnet.gr
4 Fraunhofer Institute for Applied Information Technology (FIT)
Schloss Birlinghoven, 53754 Sankt Augustin, Germany
martin.wolpers@fit.fraunhofer.de
5 European Schoolnet (EUN)
Rue de Trèves, 61, 1040 Brussels, Belgium
riina.vuorikari@eun.org

Abstract. In the world of recommender systems, it is a common practice to use public available datasets from different application environments (e.g. MovieLens, Book-Crossing, or EACMovie) in order to evaluate recommendation algorithms. These datasets are used as benchmarks to develop new recommendation algorithms and to compare them to other algorithms in given settings. In this paper, we explore datasets that capture learner interactions with tools and resources. We use the datasets to evaluate and compare the performance of different recommendation algorithms for Technology Enhanced Learning (TEL). We present an experimental comparison of the accuracy of several effective filtering algorithms applied to these TEL datasets and also implicit relevance data, such as downloads and tags, that can be augment explicit relevance evidence in order to improve the performance of recommendation algorithms.

Fig. 2. MAE of user-based, item-based and slope-one collaborative filtering

Fig. 3. F1 of user-based collaborative filtering with increasing number of neighbors
In LinkedUp we have the opportunity to apply a structured approach to develop a **community accepted evaluation framework**.

1. **Top-Down** by a literature study
2. **Bottom-up** by GCM with experts in the field
Learning Analytics and Linked Data (LALD2012)

1st International Workshop on Learning Analytics and Linked Data

Hendrik Drachsler CELSTEC NL I Abelardo Pardo University of Madrid ES

Stefan Dietze L3S, DE
Mathieu d’Aquín The Open University UK
Wolfgang Greller Open University NL
Jelena Jovanovic, University of Belgrade, SR

Wolfgang Reinhardt
University of Paderborn, DE
Katrien Verbist
K.U. Leuven, BE

#LALD

12
Development of the Evaluation Framework

P1: Initialisation
- M0-M6: Preparation
- EF proposal
- Expert validation
- Literature review
- Cognitive Mapping
- Group Concept Mapping

P2: Establishment and Evaluation
- M7-M18: Competition cycle
- Draft
- Review of EF
- New version
- Refinement of EF
- Practical experiences and refinement

P3: Exit and Sustainability
- M18-M24: Finalising
- Final release of EF
- Documentation
- Dissemination

Hendrik Drachsler 25 February 2013
Group Concept Mapping

- Group Concept Mapping resembles the Post-it notes problem solving technique and Delphi method.

- GCM involves participants in a few simple activities (generating, sorting and rating of ideas) that most people are used to.

**GCM is different in two substantial ways:**

1. Robust analysis (MDS and HCA)
   GCM takes up the original participants contribution and then quantitatively aggregate it to show their collective view (as thematic clusters).

2. Visualisation
   GCM presents the results from the analysis as conceptual maps and other graphical representations (pattern matching and go-zones).
Group Concept Mapping

brainstorm

sort

rate

...organize the issues...
Brainstorming Statements

Please generate short phrases or sentences about possible concrete indicators of the evaluation framework for evaluating Open Web Data applications in the educational domain completing the following focus statement:

"One specific indicator of the evaluation framework for assessing the Open Web Data application in the educational domain is....".

Good examples might be:
"The application enables the students to gain new insights into the topic";
"provides possibility for connecting to social network sites";
"possibility to scale up to large data sets".

Not appropriate examples are:
"technical aspects", "usability", "user scenario"

You may add as many statements as you wish. Please keep each statement brief, just one thought. Select "add this statement" after each statement or idea. Your statement will then be saved and added to the list of collected statements at the bottom of the page. You also can review the other statements to see if your idea is already there. You may search this list of collected statements using the search function below.

This task should not take longer than 30 minutes but you can always come back to add new indicators and get inspired by contributions of the other experts until Tuesday 29th of January, 12am CET.

**FOCUS PROMPT:** One specific indicator of the evaluation framework for assessing the Open Web Data application in the educational domain is....".
Group Concept Mapping

LinkedUP: Linking Web Data for Education Project

INSTRUCTIONS:
In this activity, you will categorize the statements according to their meaning or theme. Please sort each statement into a category in a way that makes sense to you. Just like grouping post-its from a brainstorming session.

1. Please read through the statements in the 'Unsorted statements' column on the left.

2. Cluster statements
Sort each statement into a category you create. Give each category a name that describes its theme or contents. You can either create categories first and then put the statements into the categories already created by drag and drop the statements into the categories, or you can start a new category by simply drag and drop a statement into the middle area. Tip: drop an item into a group when a green check mark appears.

Group the statements for how similar in meaning or theme they are to one another. There is no right or wrong way to group the statements in categories. You will probably find

Feature
- to better support group activities (both long-term course level and short-term project level)
- the student can discover new relationships between concepts which were not so obvious without the tool

Education
- It should reduce the workload without reducing quality of teaching and learning

Data
- that the used dataset(s) are sufficiently (semantically) described

Usability
- It can be easily be used by users who are not technology savvy
- general principles of usability should be considered

Hendrik Drachsler 25 February 2013
## LinkedUP: Linking Web Data for Education Project

**Priority Rating**

Rate each statement on a 1 to 5 scale for how much PRIORITY should be given to the indicator for evaluating Open Web Data applications in the educational domain, where ‘1’ equals lowest priority and ‘5’ equals highest priority.

**Project Focus Prompt:** One specific indicator of the evaluation framework for assessing the Open Web Data application in the educational domain is:...

<table>
<thead>
<tr>
<th>Rating</th>
<th>Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○ Show unrated statements only  ○ Show all statements</td>
</tr>
<tr>
<td></td>
<td>Standalone, open-source, web-based tools that can be integrated into education systems.</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 the student can discover new relationships between concepts which were not so obvious without the tool.</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 to use innovative scoring mechanism such as badges.</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 that is allows/supports knowledge exchange with other learners.</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 to build an online collaborative system to support awareness of the user (e.g. social - , task - , and/or group</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 the application is self-explanatory and therefore easy to handle.</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 the amount of data sources the tool is able to work with / connect to</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 to increase the motivation of learners to play with the data / information provided</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 that it provides knowledge to gain official certifications/ recognition</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 that the used dataset(s) are sufficiently (semantically) described</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 the comparison of peer learners or learning groups and with personal achievements</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 that it addresses a broad community of potential users</td>
</tr>
<tr>
<td></td>
<td>○ 1 ○ 2 ○ 3 ○ 4 ○ 5 it should reduce the workload without reducing quality of teaching and learning</td>
</tr>
</tbody>
</table>
• innovations in way network is delivered
• investigate corporate/structural alignment
• assist in the development of non-traditional partnerships (Rehab with the Medicine Community)
• expand investigation and knowledge of PSN/PSO's
• continue STHCS sponsored forums on public health issues (medicine managed care forum)
• inventory assets of all participating agencies (providers, Venn Diagrams)
• inventory needs of all participating agencies (providers)
• better utilization of current technological bridges
• expand and encourage utilization of interface programs to strengthen the viability and to improve the health care delivery system (teleconference)
• discussion with CCHN

...”map” the issues...
• innovations in easy network is delivered
• (investigate) corporate/structural alignment
• ongoing investigation and knowledge of non-traditional partnerships (Rehav with the Medicine Community)
• expanded investigation and knowledge of PWNS/PSO's
• continue STHCS sponsored issues on public health issues (medicine managed care focus)
• inventory needs of all participating agencies (provider, Venn Diagrams)
• sources additional funds for information expansion
• continued support by STHCS to member facilities
• expanded and encouraged utilization of interface programs to strengthen the viability and to improve the health care delivery system (teleconference)
• discussion with CCHN

...prioritize the issues...
D2.1 Evaluation Criteria and Methods

- Invited 122 external experts
- 56 experts contributed 212 indicators for the evaluation framework
- After cleaning -> 108 indicators remained
- 26 experts sorted on similarity in meaning
- 26 experts rated on priority and applicability
Look at and listen to the presentation of the Evaluation Framework

Meanwhile…create notes on

P: Plus
M: Minus
I: Interesting

Write down everything that comes to your mind, generate as many ideas as possible, do not filter your ideas.
Group Concept Mapping

A point map
Group Concept Mapping

A cluster map 15
A cluster map

1. Support Group Activities
2. Collaboration-Social
3. Privacy
4. Pedagogical Aspects
5. Assessment
6. Feedback
7. Support Learning
8. Good Searching Mechanisms
9. Performance
10. Ease of Use
11. Usability
12. Novelty
13. Data Quality
14. Data Reuse
15. Context

At Cluster 14 merged: 11 12
At Cluster 13 merged: 2 3
At Cluster 12 merged: 5 6
At Cluster 11 merged: 8 9
At Cluster 10 merged: 10 11 12
At Cluster 9 merged: 13 14
At Cluster 8 merged: 4 5 6
At Cluster 7 merged: 13 14 15
At Cluster 6 merged: 4 5 6 7

Number of Clusters:

20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2

Hendrik Drachsler
25 February 2013
Group Concept Mapping

Clusters’ labels

1. Support Group Activities
2. Privacy
3. Educational Innovation
4. Usability
5. Performance
6. Data
Rating Map Priority

Cluster Legend
Layer  Value
1  3.17 to 3.28
2  3.28 to 3.39
3  3.39 to 3.50
4  3.50 to 3.61
5  3.61 to 3.72

1. Support Group Activities
2. Privacy
3. Educational Innovation
4. Usability
5. Performance
6. Data
Rating Map Applicability

Cluster Legend

<table>
<thead>
<tr>
<th>Layer</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.84 to 2.99</td>
</tr>
<tr>
<td>2</td>
<td>2.99 to 3.15</td>
</tr>
<tr>
<td>3</td>
<td>3.15 to 3.30</td>
</tr>
<tr>
<td>4</td>
<td>3.30 to 3.46</td>
</tr>
<tr>
<td>5</td>
<td>3.46 to 3.62</td>
</tr>
</tbody>
</table>

1. Support Group Activities
2. Privacy
3. Educational Innovation
4. Usability
5. Performance
6. Data
Group Concept Mapping

Priority
Usability

Educational Innovation
Performance
Data
Privacy
Support Group Activities

Applicability
Support Group Activities
Usability
Privacy
Data
Performance
Educational Innovation

r = -0.16

Hendrik Drachsler
Group Concept Mapping

Usability

Applicability

r = -0.32

19. it can be easily be used by users who are not technology savvy
29. a fast response times also with huge amount of data
33. scalability (how it deals with increasing/large amounts of data/users/requests)
70. to be platform independent and run on multiple devices without interrupting the user experience
73. that the applications runs stable and does not
Evaluation Framework

**Educational Innovation (EI)**

- IE.1 Solves a edu. problem
- IE.2 Improves a traditional approach (More effective or efficient)
- IE.3 Provides new insights (More attractive combining data sources)
- IE.4 Collaboration patterns (e.g., connection to social networks and other user information)

**Performance (P)**

- P.1 Scalability (possibility to scale up to large datasets)
- P.2 Runs stable (Response time, amount of bugs)
- P.3 Accuracy (Information retrieval measures like Precision, Recall, F1)

**Usability (U)**

- U.1 Understandability (e.g., difficulty of logical, or navigation concept)
- U.2 Learnability (e.g., effort to get used to a new software)
- U.3 Attractive Interface

**Data (D)**

- D.1 Data coverage (Amount of used data sources, size of dataset)
- D.2 Common Data format
- D.3 Documented and version controlled
- D.4 Reliable information

**Legal (L)**

- L.1 Collects only essential data
- L.2 Secures user data
- L.3 Applies Terms-of-use
- L.4 Shows Provenance of used data
1. Literature review of suitable evaluation approaches and criteria
2. Review of comprising initiatives such as LinkedEducation, MULCE, E3FPLE and the SIG dataTEL
**Evaluation Framework**

**Educational Innovation (EI)**
- **IE.1 Improves a traditional approach**
  - Innovation in Content, Assessment practices, Organization, or infrastructure

- **IE.2 Provides new insights**
  - Learning & Teaching

- **IE.3 Collaboration patterns & Connectedness**
  - Method: Survey and expertise of judges

**Usability (U)**
- **U.1 Understandability**
  - Visibility of the system status, Match between system and the real world, Documentation

- **U.2 Learnability**
  - User control and freedom, Consistency and standards, Recognize errors

- **U.3 Attractive Interface**
  - Recognition, Efficiency of use, minimalist design
  - Method: SUS – checklist Desirability Kit

**Performance (P)**
- **P.1 Scalability**
- **P.2 Runs stable**
  - Uptime, Loading Time

- **P.3 Accuracy**
  - Precision, Recall, F1

**Data (D)**
- **D.1 Data coverage**
  - Completeness, Relevancy, Update policy, Volatility, Timeliness

- **D.2 Correctness**
  - Accuracy, Objectivity, Validity, Interlinking, Conciseness, Conformance

- **D.3 Accessibility**
  - Availability, Versatility, Understandability

- **D.4 Reliable information**
  - Provenance, Verifiability, Believability, Licensing
  - Method: Survey and report from analytic tool

**Legal (L)**
- **L.1 Collects only essential data**
- **L.2 Secures user data**
- **L.3 Applies Terms-of-use**
- **L.4 Shows Provenance of used data**

**Audience (A)**
- **A.1 User scenario**
- **A.3 Potential of the tool**
- **A.2 Reach**
  - Visitors, Page views and duration, Session Duration, Direct Connections, Ratings, Network Reach, Retweets, Social bookmarks

  - Method: Survey and report from analytic tool
Many thanks for your attention!

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