

Result Assessment Tool

Software zur Durchführung von Studien auf der Basis von Suchergebnissen

*Sebastian Sünkler, Nurce Yagci, Daniela Sygulla, Sonja
von Mach, Sebastian Schultheiß, Dirk Lewandowski*

Hochschule für Angewandte Wissenschaften Hamburg, Deutschland
{[sebastian.suenkler](mailto:sebastian.suenkler@haw-hamburg.de), [nurce.yagci](mailto:nurce.yagci@haw-hamburg.de), [daniela.sygulla](mailto:daniela.sygulla@haw-hamburg.de), [sonja.vonmach](mailto:sonja.vonmach@haw-hamburg.de),
[sebastian.schultheiss](mailto:sebastian.schultheiss@haw-hamburg.de), [dirk.lewandowski](mailto:dirk.lewandowski@haw-hamburg.de)}@haw-hamburg.de

Abstract

Das *Result Assessment Tool* (RAT) ist ein Software-Toolkit zur Durchführung von großen Studien mit Ergebnissen aus (kommerziellen) Suchmaschinen und anderen Information-Retrieval-Systemen. Die Software vereint Module, die das Design und die Verwaltung von Studien, die automatische Erfassung von Suchergebnissen durch Web Scraping sowie die Bewertung von Suchergebnissen durch Juroren und Jurorinnen anhand von Skalen in einem Bewertungs-Interface ermöglichen. Durch die Flexibilität des RATs ist es möglich, eine Vielzahl von Studientypen zu realisieren. So lassen sich neben klassischen Retrievaltests beispielsweise auch Klassifikationsstudien sowie qualitative Inhaltsanalysen auf Basis von Suchergebnissen umsetzen. Dadurch lässt sich das RAT auch fächerübergreifend einsetzen.

Schlagnworte: Suchmaschinen, Web Scraping, Retrievaltests, Software

1 Einleitung

Studien, die sich auf Ergebnisse kommerzieller Suchmaschinen und anderer Information-Retrieval-Systeme stützen, erfordern in der Regel viel manuelle Arbeit bei der Konzeption des Tests, der Erfassung der Suchergebnisse, der Suche nach Juroren und Jurorinnen, der Erfassung ihrer Bewertungen und der Analyse der Testergebnisse, was die Durchführung reproduzierbarer Stu-

dien erschwert. In der Information-Retrieval-(IR-)Community werden seit Jahren Software-Tools zu diesen Zwecken entwickelt. Dabei handelt es sich jedoch meist um Tools für die einmalige Nutzung in Studien (z. B. Bar-Ilan & Levene, 2011; Tawileh et al., 2010; Trielli & Diakopoulos, 2020), um Prototypen, die nicht weiterentwickelt wurden (Lingnau et al., 2010; Renaud & Azzopardi, 2012), sowie Software für Testkollektionen (Dussin & Ferro, 2008; Koopman, 2014; Ogilvie & Callan, 2001) oder für eng beschränkte Anwendungsfälle (The Digital Methods Initiative, 2022; Thelwall, 2009). Um alle Schritte des Testverfahrens in einer nachhaltigen Komplettlösung zu integrieren, entwickeln wir das *Result Assessment Tool* (RAT). Das RAT ist ein Toolkit, das es Forschenden ermöglicht, (groß angelegte) Studien auf der Grundlage von Ergebnissen aus Suchmaschinen und anderen Information-Retrieval-Systemen durchzuführen. Im Folgenden wird das RAT mit allen bisherigen Komponenten dargestellt.

2 Funktionsweise des *Result Assessment Tools*

Das RAT ist eine flexible webbasierte Software, die in Python mit einer PostgreSQL-Datenbank sowie Selenium als Technologie für das Web Scraping entwickelt wird. Während Forschende Zugang zu einem Interface haben, um die Studie zu entwerfen, können die Teilnehmenden ebenfalls über ein Web-Interface die Suchergebnisse zu den vordefinierten Fragen bewerten. Dabei ist das Tool so flexibel gestaltet, dass praktisch alle Studien durchführbar sind, bei denen Suchergebnisse die Grundlage bilden. Neben klassischen Retrievalstudien sind somit Klassifikationsstudien und Datenanalysen oder auch qualitative Inhaltsanalysen möglich. Insgesamt setzt sich das *Result Assessment Tool* aus fünf Modulen zusammen, die die Durchführung der Studien ermöglichen:

1. *Testgestaltung*: Hier definieren Forschende die Art der Studie, den Ergebnistyp für die Bewertung (Suchergebnisse und/oder die Trefferbeschreibungen) sowie die Art des Zugangs zum Bewertungsinterface.
2. *Suchmaschinen-Scraper*: In diesem Modul werden Suchaufgaben mit Suchanfragen definiert sowie die Suchmaschinen ausgewählt, die gescraped werden sollen. Alternativ lassen sich auch Listen mit URLs hochladen, die zur Bewertung zur Verfügung gestellt werden.

3. *Definition von Fragestellungen:* Das RAT ist sehr flexibel bei der Gestaltung von Fragen. So lassen sich neben Likert-Skalen, Schiebereglern und Fragen mit Mehrfachauswahl auch offene Fragen definieren.
4. *Bewertungs-Interface:* In dem Bewertungsinterface beantworten die Juroren und Jurorinnen die vorab definierten Fragen zu den Suchergebnissen oder Trefferbeschreibungen.
5. *Ergebnisexport:* Forschende können jederzeit die Antworten zu den Suchergebnissen als Tabellen herunterladen und für ihre Studienzwecke analysieren.

Als weiteres Modul ist eine Analysekomponente geplant, die eine Auswertung hinsichtlich gängiger Kennzahlen des Information Retrieval (MAP, NGCG usw.) sowie weitere Analysen ermöglicht.

Die Software wird nach den Kriterien der menschenzentrierten Gestaltung (User-Centered Design; UCD) entwickelt. Wesentlicher Bestandteil des UCD-Prozesses ist die Evaluation der Nutzerfreundlichkeit (Usability) der Software (Abrams et al., 2004), weshalb fortlaufende Usability-Tests sowie heuristische Evaluationen zum Einsatz kommen. Abbildung 1 zeigt die Übersichtsseite zu einer Studie im RAT. Dort werden alle Fortschritte angezeigt und es ist auf den ersten Blick erkennbar, welche Suchmaschinen, Suchanfragen, Fragestellungen und Aufgaben verwendet werden; ebenso die Anzahl der bisherigen Teilnehmenden.

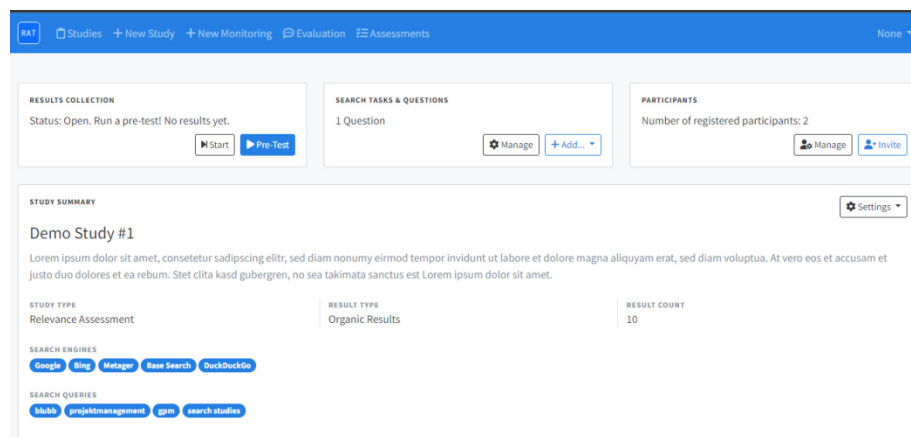


Abb. 1 Übersicht zu einer Studie im RAT

Die im RAT erzeugten Forschungsdaten (Quelltext, Ergebnisse aus durchgeführten Studien) werden von uns nach dem FAIR-Prinzip (Wilkinson

et al. 2016) über das Open Science Framework (OSF)¹ zur Verfügung gestellt. Dies ist ein Teil unserer Nachhaltigkeitsstrategie – neben dem Aufbau einer Forschenden- und Entwickelnden-Community sowie diversen Marketingmaßnahmen zum Aufbau eines internationalen Netzwerks mit Forschenden.

Danksagung

Die Entwicklung des Result Assessment Tools erfolgt im Rahmen des von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekts „Relevance Assessment Tool (RAT): Entwicklung und Erweiterung einer Software zur Durchführung von Studien, die auf Ergebnissen aus Suchsystemen beruhen“ (Förderkennzeichen 460676551; es fand eine Umbenennung des Tools statt).

Referenzen

- Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-Centered Design. In W. Bainbridge (Hrsg.), *Encyclopedia of Human-Computer Interaction* (S. 445–456). SAGE Publications.
- Bar-Ilan, J., & Levene, M. (2011). A method to assess search engine results. *Online Information Review*, 35(6), 854–868. <https://doi.org/10.1108/14684521111193166>
- Dussin, M., & Ferro, N. (2008). Design of a Digital Library System for Large-Scale Evaluation Campaigns. In B. Christensen-Dalsgaard, D. Castelli, B. Ammitzbøll Jurik, & J. Lippincott (Hrsg.), *Research and Advanced Technology for Digital Libraries* (S. 400–401). Springer. https://doi.org/10.1007/978-3-540-87599-4_45
- Koopman, B. (2014). Semantic Search as Inference. *ACM SIGIR Forum*, 48(2), Dec. 2014, 116–117. <https://doi.org/10.1145/2701583.2701601>
- Lingnau, A., Ruthven, I., Landoni, M., & van der Sluis, F. (2010). Interactive Search Interfaces for Young Children – The PuppyIR Approach. *ICALT 2010 – 10th IEEE International Conference on Advanced Learning Technologies*, 5–7 July, Sousse, Tunesia (S. 389–390). <https://doi.org/10.1109/ICALT.2010.111>
- Ogilvie, P., & Callan, J. P. (2001). Experiments Using the Lemur Toolkit. *Proceedings of The Tenth Text REtrieval Conference, TREC 2001*, Gaithersburg, Maryland, USA, Nov. 13–16, 2001.

¹ <https://osf.io/t3hg9/>

- Renaud, G., & Azzopardi, L. (2012). *SCAMP. IIX '12: Proceedings of the 4th Information Interaction in Context Symposium*, Aug. 2012 (S. 286–289). <https://doi.org/10.1145/2362724.2362776>
- Tawileh, W., Griesbaum, J., & Mandl, T. (2010). Evaluation of five web search engines in Arabic language. In M. Atzmüller, D. Benz, A. Hotho, & G. Stumme (Hrsg.), *Proceedings of LWA2010 – Lernen, Wissen, Adaptivität*, Kassel, Oct. 4–6, 2010 (8 S.).
- The Digital Methods Initiative (2022). DMI Tools. <https://wiki.digitalmethods.net/Dmi/ToolDatabase>
- Thelwall, M. (2009). *Introduction to Webometrics: Quantitative Web Research for the Social Sciences*. Morgan & Claypool. <https://doi.org/10.2200/s00176ed1v01y200903icr004>
- Trielli, D., & Diakopoulos, N. (2020). Partisan search behavior and Google results in the 2018 U.S. midterm elections. *Information, Communication & Society*, 25(1), 145–161. <https://doi.org/10.1080/1369118X.2020.1764605>
- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, Art. No. 160018. <https://doi.org/10.1038/sdata.2016.18>

In: A. Imeri, K. Scheibe, F. Zimmer (Hrsg.): *Informationswissenschaft im Wandel. Wissenschaftliche Tagung 2022 (IWWT22)*. Düsseldorfer Konferenz der Informationswissenschaft, 6.–7. Oktober 2022, Haus der Universität Düsseldorf. Glückstadt: Verlag Werner Hülsbusch, S. 282–286. DOI: <https://doi.org/10.5281/zenodo.7457892>.

Analyse von kritischen Erfolgsfaktoren bei der Entwicklung von Offenen Bildungsressourcen anhand von Fallstudien

Stefan Dreisiebner, Lea Wöbbekind, Thomas Mandl

Universität Hildesheim, Deutschland

{dreisiebner, woebbek, mandl}@uni-hildesheim.de

Abstract

Das Poster steht im Kontext des EU-geförderten Projekts *Digital Education for Crisis Situations* (DECriS), welches vor dem Hintergrund der Corona-Pandemie gestartet wurde und sich insbesondere mit dem Einsatz von Offenen Bildungsressourcen (OER) beschäftigt. Der Fokus dieses Beitrags liegt auf der Analyse von kritischen Erfolgsfaktoren bei der Entwicklung von OER anhand von Fallstudien. Dieses Teilprojekt verfolgt das Ziel, in einem früheren Teilprojekt definierte kritische Erfolgsfaktoren für die Evaluierung von OER auf deren Praxisrelevanz zu untersuchen. Dies erfolgt mittels eines mehrstufigen Ansatzes, bei dem einerseits Studierende mittels Fragebögen befragt und andererseits Lehrende mittels teilstrukturierten Interviews interviewt werden. Daraus sollen mehrere Fallstudien entstehen, die beispielhaft aufzeigen, wie die aufgestellten Erfolgsfaktoren angewendet werden könnten. Das Poster liefert einen Einblick in erste Ergebnisse aus der Verifizierung der Erfolgsfaktoren unter Studierenden, basierend auf einer Online-Befragung an den Projekt-Partnerinstitutionen im September 2022. Hierbei zeigt sich insbesondere, wie Studierende die Erfolgsfaktoren nach deren Relevanz reihen würden, sowie deren eigene Erfahrungen mit OER und Erwartungshaltungen an OER.

Schlagworte: Fallstudie, OER, Open Educational Resources, Erfolgsfaktoren, Evaluierung

1 Problemaufriss und Projektziele

Unter Offenen Bildungsressourcen (OER) können Lehr-, Lern- und Forschungsmaterialien verstanden werden, welche einer offenen Lizenzierung unterliegen, um deren freie Wiederverwendung, laufende Verbesserung und Adaptierung durch Dritte zu Bildungszwecken zu ermöglichen (Orr et al., 2015). OER haben hierdurch auch das Potenzial, einen gesellschaftlichen Beitrag zur Lösung von Herausforderungen im Bildungsbereich zu leisten, speziell zur Fortbildung von Lehrenden, Reduzierung von Bildungskosten, Verbesserung der Qualität von Lernressourcen sowie der leichteren Zugänglichkeit und größeren Verbreitung von hoch qualitativen Lernressourcen (Orr et al., 2015).

Neben offen lizenzierten Online-Kursen (Dreisiebner et al., 2021) und kommerziellen Weiterbildungs-Anbietern wie Coursera¹ entstehen zunehmend auch OER-Repositoryen wie twillo² mit dem Ziel, OER besser zugänglich zu machen. Dennoch bleibt die korrekte Lizenzierung von OER neben der Auffindbarkeit geeigneter OER weiterhin eine Herausforderung für Lehrende und Studierende (Anderson & Leachman, 2019). Speziell im Kontext von Online-Lehre zeigten sich im Rahmen der Corona-Krise auch mangelnde technische Kenntnisse als Barriere (Boté-Vericad, 2021), zugleich aber auch das in diesem Kontext empfundene Gefühl der Herausforderung als motivierender Faktor (Feldhammer-Kahr et al., 2021). Um die Verbreitung und Nutzung von OER zu erhöhen, bedarf es tiefergehender Analysen der hemmenden Faktoren (Mandl, 2021).

Vor diesem Hintergrund startete das EU-geförderte Projekt *Digital Education for Crisis Situations* (DECriS, <https://decris.ffos.hr>). Ausgangspunkt des DECriS-Projekts war eine Analyse des Einsatzes von Online-Lehre und OER während der Corona-Krise in den Studienjahren 2019/20 und 2020/21 an informations- und bibliothekswissenschaftlichen Hochschulinstitutionen in Europa, basierend auf einer Karte der europäischen informations- und bibliothekswissenschaftlichen Bildungslandschaft (Meschede et al., 2018).

1 <https://www.coursera.org/>

2 <https://www.twillo.de>

2 Methodik, erste Ergebnisse und Ausblick

Dieser Beitrag fokussiert sich auf das Projektziel „Analyse von kritischen Erfolgsfaktoren bei der Entwicklung von OER anhand von Fallstudien“, welches wiederum auf vorherigen Projektergebnissen aufbaut. Wesentliche Grundlage hierfür ist eine strukturierte Literaturanalyse von bestehenden Publikationen zur Evaluierung von OER (Todorova et al., 2022). Diese basierte auf einer Abfrage in Web of Science, SCOPUS, Google Scholar und bei relevanten internationalen Organisationen wie der UNESCO zu Anfang des Jahres 2022. Aus den 199 Ergebnissen wurden anhand von vorab definierten Inklusions-Kriterien 71 relevante Dokumente identifiziert, welche in einem nächsten Schritt im Volltext einer qualitativen Inhaltsanalyse unterzogen wurden. Hieraus konnten schlussendlich folgende kritische Erfolgsfaktoren für die Evaluierung von OER abgeleitet werden (Todorova et al., 2022):

- Lerninhalte und Gestaltung der Lernerfahrung
- Lernmanagementsystem (LMS) und technische Tools
- Lernprozess und Pädagogik
- Mehrwert stiftende Zusatzdienste.

Diese kritischen Erfolgsfaktoren wurden jeweils durch untergeordnete Sub-Kriterien genauer spezifiziert. Zwei Beispiele solcher Sub-Kriterien für „Lerninhalte und Gestaltung der Lernerfahrung“ sind inhaltliche Qualität und Interaktivität. Deren Praxisrelevanz soll in einem nächsten Schritt mittels eines mehrstufigen Ansatzes analysiert werden: Einerseits sollen Studierende nach deren Relevanzeinschätzung befragt werden, indem die Erfolgsfaktoren unter anderem gereiht werden. Andererseits werden aus den vorhergehenden Erhebungen besonders im OER-Kontext engagierte Lehrende identifiziert und im Zuge von teilstrukturierten Interviews erhoben, inwiefern diese Erfolgsfaktoren in deren Arbeit bereits angewendet werden oder potenziell anwendbar wären. Daraus sollen mehrere Fallstudien entstehen, die beispielhaft aufzeigen, wie die aufgestellten Erfolgsfaktoren angewendet werden könnten.

Das Poster wird neben einem Überblick über das Gesamtprojekt einen Einblick in erste Ergebnisse aus der Verifizierung der Erfolgsfaktoren unter Studierenden geben, basierend auf einer mittels LimeSurvey implementierten Online-Befragung. Der Fragebogen enthält Fragen zu dem Verständnis von OER, persönlicher Erfahrung mit OER und deren Einsatzbereichen, Gründen für Nutzung und Nicht-Nutzung, der persönlichen Sicht auf die Zukunft von

OER sowie den persönlichen Erwartungshaltungen an OER. Ebenso sollen die im Projekt definierten Erfolgsfaktoren nach der empfundenen Relevanz gereiht werden. Eine offene Frage erlaubt die Ergänzung weiterer Erfolgsfaktoren. Die Umfrage wurde innerhalb der Projekt-Partnerinstitutionen im September 2022 per E-Mail unter Studierenden verbreitet. Erste Ergebnisse werden mittels Microsoft Excel statistisch und grafisch aufbereitet. Das Endergebnis soll zu einem besseren Verständnis der kritischen Erfolgsfaktoren für die OER-Nutzung, speziell während der Corona-Krise, führen. Hierbei soll insbesondere auch aufgezeigt werden, wie weit die digitale Transformation in der Lehre schon fortgeschritten ist und in welchen Bereichen noch Barrieren bestehen.

Danksagung

Dieses Projekt wurde mit Unterstützung der Europäischen Kommission finanziert. Die Verantwortung für den Inhalt dieser Veröffentlichung tragen allein die Verfasser; die Kommission haftet nicht für die weitere Verwendung der darin enthaltenen Angaben.

Referenzen

- Anderson, T., & Leachman, C. (2019). Strategies for Supporting OER Adoption through Faculty and Instructor Use of a Federated Search Tool. *Journal of Librarianship and Scholarly Communication*, 7(1). <https://doi.org/10.7710/2162-3309.2279>
- Boté-Vericad, J.-J. (2021). Perceived barriers for distance teaching in higher education during the COVID-19 crisis: “I never did a video before”. *Education for Information*, 37(3), 377–397. <https://doi.org/10.3233/EFI-200418>
- Dreisiebner, S., Polzer, A. K., Robinson, L., Libbrecht, P., Boté-Vericad, J.-J., Urbano, C., Mandl, T., Vilar, P., Žumer, M., Juric, M., Pehar, F., & Stričević, I. (2021). Facilitation of information literacy through a multilingual MOOC considering cultural aspects. *Journal of Documentation*, 77(3), 777–797. <https://doi.org/10.1108/JD-06-2020-0099>
- Feldhammer-Kahr, M., Tulis, M., Leen-Thomele, E., Dreisiebner, S., Macher, D., Arendasy, M., & Paechter, M. (2021). It's a Challenge, Not a Threat: Lecturers' Satisfaction During the Covid-19 Summer Semester of 2020. *Frontiers in Psychology*, 12, 638898. <https://doi.org/10.3389/fpsyg.2021.638898>

- Mandl, T. (2021). The Covid Pandemic as Crisis and Catalyst: Future Roles, Skills and Tasks of Information Professionals. *Proceedings of the 4th I-LISS International Conference – IIC 2021* (S. 283–293).
- Meschede, C., Ortiz-Repiso, V., & Kluin, M. (2018). Library and Information Science education in Europe: Building an interactive map. *Proceedings of the International Symposium on the Future of Education in Information Science (FEIS)*, 10–11 Sept. 2018, Pisa, Italy (S. 217–222).
- Orr, D., Rimini, M., & Van Damme, D. (2015). *Open Educational Resources. A catalyst for innovation*. OECD. <https://doi.org/10.1787/9789264247543-en>
- Todorova, T., Pavlova, D., Kovatcheva, E., & Bogova, H. (2022): Critical factors for evaluation of existing OER's: Research in progress [Presentation], Hybrid Multiplier Event Erasmus+ Project DECriS: Critical success factors and their typological classification for the evaluation of the Open Educational Resources (OER), Sofia, Bulgaria.