iConference 2025 (ID: 126)

Topic Mining and Evolution Analysis of Digital Literacy Research Topics Abroad

Zheng Li[1], Tao Zhang[1]

[1]School of Information Management, Heilongjiang University, Harbin 150080, China

Abstract

The study selected literature data on digital literacy from the Web of Science database between 2004 and 2023 as research subjects, dividing them into ten stages. The paper employs text analysis techniques to conduct a systematic quantitative analysis of literature in the field of digital literacy, aiming to recognize and organize the research topics and their evolutionary process within this field.

Introduction

With the rapid development of new-generation information technologies such as mobile internet and cloud computing, coupled with the deepening global digital transformation, there is a continuous increase in the demand for effective retrieval, evaluation, utilization, and creation of digital information. Digital literacy has become a crucial capability for citizens to participate in modern social economic and cultural activities. Concurrently, scholars' research on the field of digital literacy is also constantly transforming and updating. To better address these changes, it is crucial to gain a deep understanding of the research hotspots and evolutionary processes in the field of digital literacy. In recent years, with the advancement of machine learning and text mining methods, it has become popular to use topic models to detect knowledge structures from the perspective of specific subject fields.

Methodology

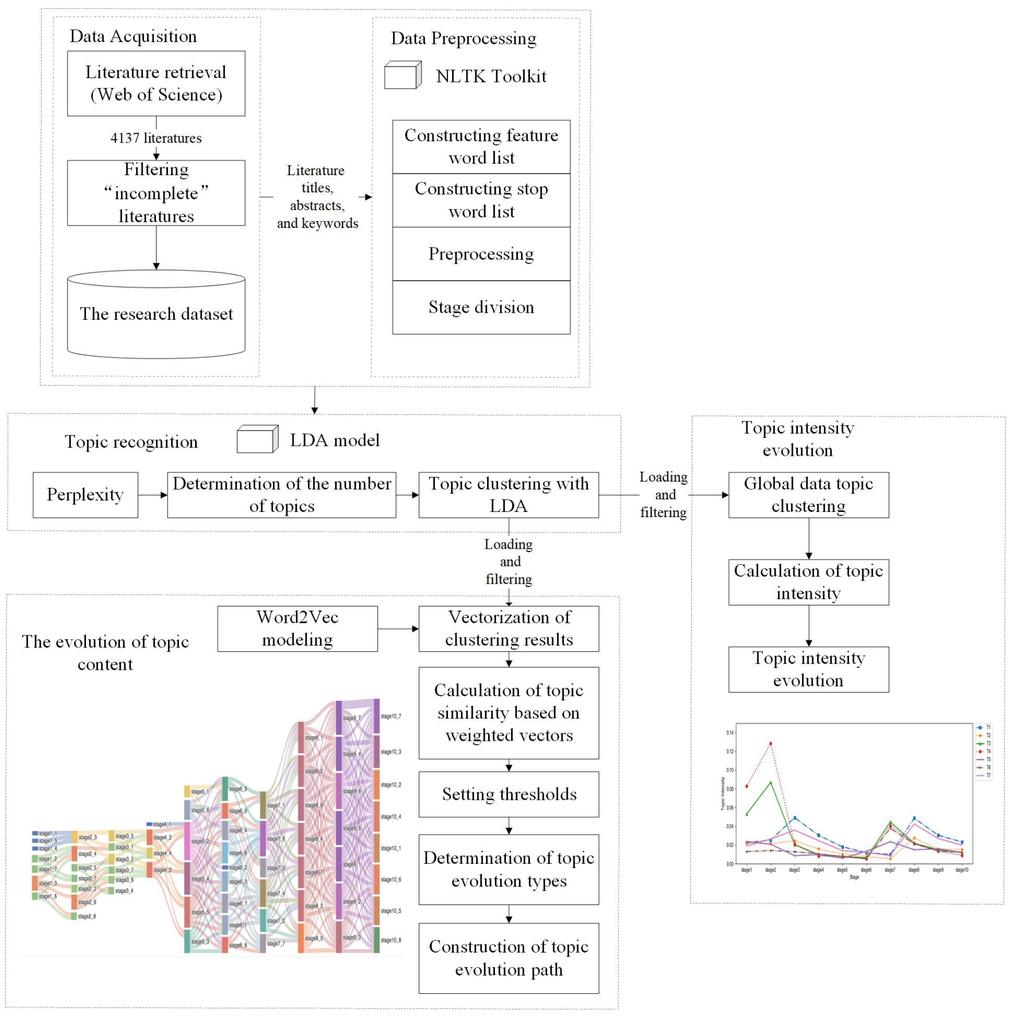

Fig1. Overall analysis framework diagram.

1. Data acquisition and preprocessing. We retrieved literature data from the Web of Science database and performed text preprocessing.
2. Topic recognition. We used the LDA model to cluster the literature and identify research topics for each stage.
3. The evolution of topic content. By calculating the similarity between topics in adjacent stages, we analyzed the evolutionary paths of topic content.
4. Topic intensity evolution. We calculated the intensity of each topic across different stages to analyze its evolutionary trends.

Results

Tab1. Stage topic words distribution.

Stages	Topic words
Stage 1	1_1(Knowledge Society); 1_2(ICT Education); 1_3(Survey Measures of Digital Literacy); 1_4(Digital Literacy Project); 1_5(Multiple Literacy); 1_6(Media Impact); 1_7(Visual Literacy)
Stage 2	2_1(Education in the Information Society); 2_2(Digital Divide); 2_3(Second Language Digital Literacy); 2_4(Capacity for Lifelong Learning); 2_5(Practice Project); 2_6(Online Teaching); 2_7(Digital Literacy in the European Union); 2_8(Computer Network User Survey)
Stage 3	3_1(Children's Literacy); 3_2(Social Media); 3_3(Print Text); 3_4(Gender Differences in Information Literacy among University Students); 3_5(E-service); 3_6(Computer-Based Second Language Literacy); 3_7(Literacy Education)
Stage 4	4_1(ICT and Education); 4_2(Network Based Computer Education); 4_3(Applications for Mobile Tools); 4_4(Computer Literacy); 4_5(Digital Reading); 4_6(Game-Based Learning)
Stage 5	5_1(Digital Knowledge Practice Age); 5_2(Digital Game-Based Learning); 5_3(ICT Curriculum); 5_4(Models of Digital Practice); 5_5(Digital Literacy Practice); 5_6(Development of Language Education); 5_7(Education Innovation)
Stage 6	6_1(Online Courses); 6_2(ICT in Education and Social Development); 6_3(Online Learning); 6_4(Information Technology Education for Children); 6_5(Online Social Networks); 6_6(Information Privacy); 6_7(Web 2.0 Technology); 6_8(Online Information Content); 6_9(Network Education)
Stage 7	7_1(Digital Literacy Framework); 7_2(Online Education Tools); 7_3(Educational Practices); 7_4(Digital Competence); 7_5(Library Information Service); 7_6(Health Information on the Internet); 7_7(Library and Digital skills)
Stage 8	8_1(Digital Literacy Education); 8_2(E-Health Information Literacy); 8_3(Education and Development in Digital Technology); 8_4(Online Learning Community); 8_5(Online Information Education); 8_6(Digital Technology Practices for Children); 8_7(Social Media Education)
Stage 9	9_1(Technology Socialization); 9_2(Education and Practice of Digital Technology); 9_3(COVID-19th Era of Online Education); 9_4(Digital Technology for Older Adults); 9_5(Social Media Literacy in the COVID-19 Era); 9_6(Children's Critical Digital Literacy); 9_7(Professional Development and Challenges)
Stage 10	10_1(Online Service); 10_2(Health Service Challenges); 10-3(Analysis and Identification of Children's Digital Literacy); 10-4(Digital Transformation); 10-5(Development and Analysis of Digital Competencies); 10-6(Health Problems of Older Adults in Society); 10-7(Factors Impacting Education); 10-8(Challenges of Online Education)

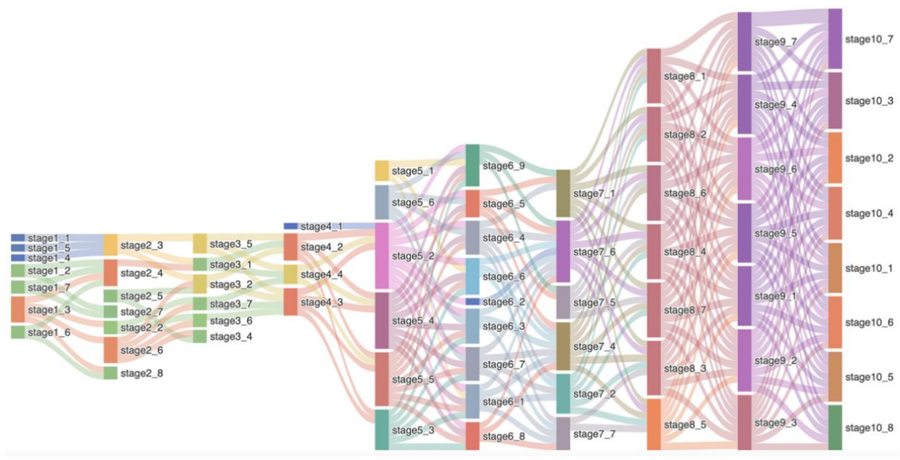

Fig2. The topic content evolutionary path.

Tab2. Topic distribution on the global stage.

Topics	Topic words(Top 10)
T1 Children's Education	children, development, education, knowledge, challenge,practice, literacy, focus, article, curriculum
T2 Social Development	government, the_network, social, community, citizen,age, ict, information, governance, development
T3 Digital Technology	service, learner, tool, development, benefit,telehealth, technology, context, privacy, challenge
T4 Educational Development	development, education, payment, national, project,financial, implementation, informatics, crowdsourcing, university
T5 Health	health, group, factor, efficacy, children,digital_competence, older_adult, digital_technology, risk, background
T6 Library Information Resources	information, resource, library, librarian, role,quality, evaluation, model, factor, online
T7 Online Education	online, mooc, challenge, teach, information, content, tool, education, language, the_network

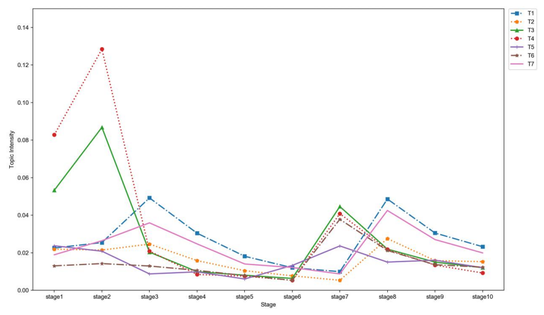

Fig3. Evolutionary trend of topic of intensity.

Discussion

- Topic Evolution: Exploration, Practice, and Expansion.
- Intensity Evolution: The Combined Effect of Polocies, Research Content, and Main Field.
- Future Research Topics: Education, Technology, and Social Topics.

Conclusions

Through the LDA2vec model, we have successfully revealed the research hotspots and their evolutionary trends in the field of digital literacy. We predict that in the future, the topic of education and digital technology will remain the focus of research in the field of digital literacy. At the same time, the field will concentrate on conducting in-depth research into social topics such as health, aging society, and the challenges and risks associated with digital transformation.