Research on Construction Methods for Scholar Databases in Digital Libraries: Postprint

Authors: Zheng Ang, Jianxun Zeng

Date: 2023-04-01T16:15:50+00:00

Abstract

[Purpose/Significance] From the perspective of digital library resource utilization and organization, this paper designs approaches for scholar data identification and scholar database construction to help improve the efficiency of digital library resource construction and specialized services. [Method/Process] This study investigates the research and practice of scholar databases domestically and internationally from four aspects: scholar selection and inclusion sources, scholar description content and its framework, scholar database construction, and scholar database application methods. By analyzing scholar characteristic attributes and studying structured expression methods for scholar data, it proposes a construction process and overall framework for scholar databases based on digital libraries. [Results/Conclusion] It proposes a parallel advancement strategy for scholar database construction and application, emphasizing that scholar databases should be integrated into the scientific research management process, mobilize scholars to participate in construction, enhance display and publicity effects, combine with talent identification, serve team and thematic resource construction; combine with knowledge management, balance the functions of scholar archiving and scholar profiling, and expand precision service functions.

Full Text

Preamble

ChinaXiv Cooperative Journal
Vol. 64, No. 5, March 2020

Research on the Construction Method of Scholar Repository Based on Digital Libraries

Zheng Ang, Zeng Jianxun
Institute of Scientific and Technical Information of China, Beijing 100038

Abstract:
[Purpose/Significance] From the perspective of digital library resource utilization and organization, this paper designs a method for scholar data identification and scholar repository construction to enhance the efficiency and distinctive services of digital library resource development. [Method/Process] The study investigates domestic and international research and practices on scholar repositories across four dimensions: scholar selection and collection sources, scholar description content and framework, repository construction methods, and application approaches. By analyzing scholar characteristic attributes and studying structured representation methods for scholar data, the paper proposes a construction process and overall framework for scholar repositories based on digital libraries. [Results/Conclusion] The paper puts forward a synchronized promotion strategy for scholar repository construction and application, emphasizing that repositories should be integrated into research management processes, mobilize scholars to participate in construction, enhance exhibition and publicity effects, combine with talent identification to serve team and thematic resource development, and integrate with knowledge management to fulfill both scholar archiving and profiling functions, thereby expanding precision service capabilities.

Keywords: scholar repository; digital library; institutional repository; scholar identification
Classification Number: G250.74
DOI: 10.13266/j.issn.0252-3116.2020.05.014

Scholar repositories organize resources around scholars to describe their academic characteristics, serving as data systems for storing, retrieving, utilizing, and discovering scholarly research outputs. These repositories not only catalog and identify scholars' academic features, attributes, and achievements, but also describe and link their academic relationships, career trajectories, and professional paths. The resulting academic resource collections constitute fundamental units for building institutional repositories, basic elements for evaluating scholarly performance, and essential materials for demonstrating institutional strength and individual scholar accomplishments [1]. Consequently, scholar repository construction represents not only a critical topic in the development of distinctive resources for digital libraries and research platforms, but also an important measure for libraries and research organizations to provide precise services for research evaluation and scientific personnel.

In recent years, database vendors, universities, research institutions, and some research funding agencies have developed overall frameworks and construction processes for digital library scholar repositories based on self-built resources and

platforms, employing multi-source data integration to optimize foundational resources and proposing advancement strategies for repository construction and application. However, relying solely on commercial databases or institutional repository resources, with their exclusive focus on academic output integration and measurement, fails to comprehensively reveal scholar characteristics or fully cover scholarly achievements. To address this limitation, this paper aims to optimize the construction process based on digital libraries by improving the scholar repository metadata system.

1. Current Status of Scholar Repository Construction

Scholar repository construction primarily involves revealing scholars' research activities, communication behaviors, academic relationships, and output achievements, as well as scholar evaluation, identification, and services. This study analyzes the collection scope and selection methods of domestic and international scholar repositories, investigates scholar description systems and frameworks, and examines construction methods and application status to identify essential steps in repository development. Scholar selection methods and collection sources affect construction effectiveness, while feature description constitutes a key component and prerequisite for effective application.

1.1 Scholar Selection and Collection Sources

Different construction objectives lead to varying selection scopes and resource acquisition approaches. Commercial databases and academic search engines establish repositories for scholars with research outputs based on specific screening criteria. AMiner targets experts in artificial intelligence and related fields, integrating relevant papers through large-scale computation to identify target scholars. Baidu Academic automatically aggregates academic achievements for scholars with certain publication and citation thresholds, while also allowing other scholars to claim their outputs and build personal homepages, having generated over 4 million scholar pages to date. Although commercial databases and search engines have constructed numerous scholar pages through automatic aggregation, scholar claims remain limited—for instance, the CNKI Scholar Repository encompasses 12 million scholars but only 100,000 have claimed their achievement information [2]. Universities and research institutions typically restrict selection to their own faculty, researchers, graduate students, undergraduates, and staff, as exemplified by Xi'an Jiaotong University's XJTU Academic Hub [3].

The scope of academic achievement collection significantly impacts repository effectiveness. Database vendors usually rely on their own 收录 resources, such as CNKI Scholar Repository's foundation in the CNKI Chinese Journal Full-text Database. This approach is limited by resource scope and cannot comprehensively reveal scholars' achievements or encompass non-academic information. Integrating multi-source data yields richer, more complete scholarly outputs. Baidu Academic obtains bibliographic data through partnerships with content

providers, harvests open resources via the OAI-PMH protocol, and crawls data through search engines to compile scholars' Chinese and international academic achievements. Universities and institutions often base construction on purchased and self-built digital academic resources—Tsinghua University and Xi'an Jiaotong University, for example, use databases such as WoS, EI, Nature, and Science as collection scope combined with institutional repository resources [3], supplemented by scholars' personal information submissions. While this approach yields satisfactory initial results, subsequent maintenance and updates struggle to ensure timeliness and accuracy.

1.2 Scholar Description Content and Framework

Repository construction requires organizing and describing scholar characteristics, achievements, and relationships to enable presentation and application. Data vendors, research institutions, and scholar identifier systems integrate scholar data through database consolidation, web crawling, and research output registration, resulting in varying description content and effectiveness. Unique identifier systems like ResearcherID and ORCID target global scholars, maximizing the display of citation and co-authorship relationships [4]. University and institutional repositories reveal publication and citation data comprehensively and timely, integrating academic experience, publication distributions, disciplinary topics, and collaborator information [5], while leveraging database and knowledge base advantages to provide full texts or links.

To enhance semantic information extraction and description, discovery service search systems construct storable and computable scholar description frameworks that express scholars and related research entities and relationships, offering references for optimizing digital library scholar description methods. AMiner establishes a scholar description ontology by extending the FOAF framework, defining entities including researchers and publications with 24 attributes and two relationship pairs (collaborators and creation), enabling better reasoning and mining of relationships between academic entities to derive more diverse characteristic indicators such as social influence and activity levels [6]. To fuse the billion-scale heterogeneous data of Microsoft Academic Graph (MAG) and AMiner Academic Graph, the Open Academic Graph (OAG) establishes venue, author, and paper schemas with entity and attribute frameworks, creating 65 million matching relationships to structurally describe publishers, papers, and authors [7].

1.3 Scholar Repository Construction Methods

Most current repositories combine automation with crowdsourcing, automatically constructing repositories from digital library literature resources and subsequently encouraging scholars to manually review and improve information through various incentives. The key to automated construction lies in linking academic achievements to scholars, inevitably encountering name disambiguation issues that require distinguishing identities and resources of scholars with

identical names—a major challenge in current research and practice. To accurately locate scholars and their outputs in massive academic resources, AMiner employs network analysis methods based on entity relationship weights to analyze ego-network characteristics of name-sharing scholars and their belonging to different clusters, achieving disambiguation through community division [8]. CNKI and Wanfang primarily use "name + affiliation" combinations for disambiguation [9-10]. Tsinghua University Scholar Repository selects valuable scholars to establish unique identifiers (THUID), initiates automatic publication tracking projects, and formulates complete analysis and tracking strategies [11]. Other research explores disambiguation methods for author names [12-13] or attempts scholar identification by linking ORCID, ResearcherID, and other unique identifiers with establishing authority files [14].

Regarding manual review and improvement, current institutional repositories employ research management measures and resource access incentives to promote scholar participation in research output registration. Xiamen University integrates its repository with research information management platforms, sharing data with unified identity authentication systems to supplement and correct repository data based on scholar feedback [15]. ResearchGate requires registration for resource access, matching scholars' self-registration with existing academic resources, allowing submission of documents, links, or proofs that undergo review before completing registration. Theoretical research has also designed scholar identification, matching, push-claim, and supplementary claim workflows based on institutional and scholar repositories [16-17].

1.4 Scholar Repository Application Methods

Most repositories provide scholar search pages and personal homepages displaying basic information, research achievements, and updates. The University of Macau Scholar Repository offers 14 search fields including ORCID, title, and author, supporting image, advanced, and professional searches [18]. AMiner serves as a data provider for Sogou Academic Search, increasing usage frequency [19]. Tsinghua University, Lanzhou University, and University of Macau homepages highlight their scholars' publications in top-tier journals like Cell, Nature, and Science, featuring columns for "highly-cited/hot papers" and "this issue's recommendations" to regularly promote popular articles and scholars [20].

Xiamen University links its repository with research output, using it as foundational data for annual performance evaluation, professional title assessment, and project application and management, establishing an independent evaluation indicator database and employing visualization tools to provide decision support for university management [15]. Since 2017, Tsinghua University's repository has served as the data source for professional title applications and annual faculty assessments [11]. Beyond serving scholars and research departments, repositories are utilized in talent mining. AMiner provides intelligent services by building expert profile systems for the National Natural Science Foundation and expert portrait databases for the Ministry of Science and Technology, while also

establishing talent maps for Alibaba and expert systems for CFF to serve enterprises and research institutions. ResearchGate calculates institutional research levels through scholar-institution associations, helping scholars quickly identify collaborative projects, institutions, and scholars, and providing recruitment services that enable institutions and individuals to hire high-quality researchers [21].

In summary, scholar repositories have developed rapidly in recent years, with name disambiguation, automatic publication tracking, and unique identifier establishment becoming common technical methods. Artificial intelligence and machine learning have begun applying to repository construction through semantic mining and deep learning to establish ontologies or structured description systems. However, challenges remain: (1) Incomplete feature revelation, with overemphasis on output integration and measurement while lacking in-depth reasoning of academic relationships and entity feature mining; (2) Single data sources primarily based on digital library resources without fusion of massive web resources, failing to leverage distributed resource management advantages for comprehensive integration; (3) Insufficient application drivers, limited to basic functions like page generation and search, without maximizing effectiveness in knowledge archiving, trajectory display, or integration with research management and evaluation.

2. Scholar Characteristics and Metadata Model

Digital library-based scholar repositories must reflect scholars' basic information for disambiguation in literature services while deeply revealing academic attributes to provide data foundations for personalized services. Repositories should effectively organize metadata reflecting scholar attribute features, extracting and identifying elements from massive academic resources based on application objectives and requirements to form structured description frameworks. This requires orderly organization of scholar information, identification and revelation of academic attributes, and accurate grasp of scholar characteristics to dynamically reflect academic trajectories.

2.1 Scholar Characteristic Attribute Analysis

Scholars are professionals engaged in research in scientific, cultural, and educational fields [22], possessing characteristic entities and attributes including professional education, academic degrees, affiliation types (research institutes, universities, corporate R&D departments), research fields, disciplinary expertise, publications, patents, academic honors, and academic networks. Individual scholars exhibit different features based on their academic experiences and disciplines—for instance, humanities and social sciences scholars rarely hold patents. These attributes are dispersed across digital library registration information, literature databases, personal webpages, academic news, and social networks, reflecting diverse scholar characteristics. Therefore, scholar feature

selection should address digital library repository application needs, considering scenarios such as scholar identification, research evaluation, talent mining, and personalized services, while ensuring comprehensive revelation of academic features to design a framework that accurately reflects both common academic characteristics and individual scholar traits.

Through investigation of literature [23-25] and repositories including AMiner, CNKI, Baidu Academic, Tsinghua University, and Peking University, and considering data sources, scholar features, and application scenarios, this study constructs a "Scholar Dimension-Element" characteristic attribute framework (Figure 1 [Figure 1: see original paper]). Since scholar attribute features appear with varying frequencies, the framework should allow for partial repetition or absence of features. Regular expression rules express element occurrence frequencies: "*" indicates zero or multiple occurrences; "?" indicates zero or one occurrence; "+" indicates one or multiple occurrences; no symbol indicates exactly one required occurrence. The proposed framework includes six dimensions comprising 27 elements: basic information reflects natural attributes; communication information facilitates academic exchange and serves as foundational data for name authority and scholar identification; academic orientation reflects research directions and expertise for precision research services; academic relationships include collaborating scholars and institutions from formal and informal exchanges, reflecting networks and activity levels; honors, papers, patents, monographs, and funded projects reflect academic achievements; education and work experiences reflect academic backgrounds and expertise, forming the basis for evaluation and talent mining.

2.2 Scholar Metadata Model

Repository construction involves not only matching scholars with literature data but also research entities including achievements and institutions, with logical relationships existing between different entities and attributes. Digital libraries can therefore adopt entity-relationship network methods to achieve reasoning and mining through links between research entities. Based on the scholar characteristic attributes in Figure 1, this study designs a digital library scholar metadata model (Figure 2 [Figure 2: see original paper]) using entity-relationship-attribute expressions to achieve structured representation and dynamic association. Papers, honors, and other achievements and institutions are converted into entities with expanded attributes. Academic orientation, which cannot be converted to an entity, is directly associated with the "scholar" entity through discipline and research direction attributes. Collaborators and institutions in academic relationships can be realized through author and affiliation attributes in paper and patent entities. Some scholar attributes result from combinations of scholars and research entities and cannot be assigned to scholars or other entities alone, thus belonging to relationships between entities. For example, degrees, graduation dates, and majors belong to scholars' educational experiences rather than inherent institutional attributes, as scholars acquire these

attributes only when associated with specific institutions—therefore, these attributes belong to the "education" relationship.

To enable association and storage of multi-source data, logical structure design of scholar metadata is necessary for relational database construction. Following the Third Normal Form (3NF), the scholar metadata E-R model is converted to a relational model satisfying first and second normal forms, creating related data tables to enable inter-table associations (Figure 3 [Figure 3: see original paper]). Scholar ID links to scholar achievement information tables for centralized display of all outputs and serves as the foundation for linking scholar-related attributes or features, enabling corresponding updates to related tables and fields when scholar data changes.

3. Overall Framework and Construction Process Based on Digital Library

3.1 Overall Framework Design

Constructing scholar repositories based on digital libraries requires leveraging digital library technical and information architecture, utilizing resource processing and collection systems, heterogeneous resource integration systems, digital resource management systems, resource scheduling systems, and user management platforms to design an overall framework (Figure 4 [Figure 4: see original paper]). The framework collects and processes scholar-related academic data and digital objects from diverse sources, conducts academic network modeling and analysis, and ultimately achieves scholar data application.

Digital library scholar repositories use internet resources and digital library resources as data sources. Through digital resource collection and processing systems, they harvest academic resource metadata based on the OAI-PMH protocol and collect digitized documents and publications. The digital object system describes resources according to digital object standards and rules to generate metadata and 调度 codes that constitute digital objects. The integration layer performs deduplication and standardization processing, while heterogeneous resource integration systems enable consolidation of internal and external metadata and resources. Based on digital resource management and storage systems and considering distributed storage characteristics and multi-source, multi-channel distribution of academic resources, repositories adopt centralized metadata storage with distributed digital object storage. On this foundation, scholar modeling and analysis integrate data according to the scholar metadata framework to form scholar tag systems for profiling. Mining and analysis based on literature data and social networks construct academic network models from 网状 associations between scholars and resources to reveal collaboration networks. Scholar clustering based on characteristics identifies similar scholars and reveals academic teams. Finally, service platforms built on digital library resource publishing and user retrieval systems apply scholar resources to evaluation, profiling, knowledge management, research management, retrieval, and

precision recommendation.

3.2 Construction Process

The construction process relies on digital library resources and platforms to collect data from multiple sources, achieve scholar identification through name authority files and unique identifiers, aggregate, disambiguate, and clean data sources to form basic scholar resource sets. Based on selection strategies, target scholars are identified and information confirmed through scholar claims. Feature mining and relationship extraction complete data processing, ultimately enabling repository services and applications. The construction flow is illustrated in Figure 5 [Figure 5: see original paper], with key steps as follows:

3.2.1 Multi-source Data Collection

Digital libraries should leverage distributed resource management characteristics to cooperate with renowned domestic and international database vendors, enriching foundational academic literature resources. Employing machine learning principles and automatic tracking, they should mine scholar achievement and usage information from academic literature resources while discovering and collecting scholar-related webpages including personal homepages, biographical entries, and academic news to obtain latest academic developments. Repository construction requires establishing update mechanisms for continuous resource collection and monitoring webpage layouts and metadata changes to timely address data collection issues.

3.2.2 Scholar Data Integration

Cleaning, integrating, and aggregating multi-source academic information around scholars constitutes critical repository development work. Data cleaning achieves standardization, eliminates low-quality data, and supplements missing fields. Data integration converges different sources and handles conflicts in partially inconsistent fields. Using unique identifiers such as ORCID and ResearcherID, name authority files, and machine learning techniques, repositories perform name disambiguation and link resources to scholars for precise, comprehensive resource sets.

3.2.3 Scholar Selection

Using authors from digital library 收录 literature as the selection scope, repositories must establish selection criteria based on application objectives, setting quantitative thresholds (publication volume, citations) or qualitative indicators (scholar status). Selection strategies should identify valuable scholars from perspectives of identity, achievements, and expertise, such as high-productivity/high-citation authors, academicians, "Thousand Talents Program" scholars, or discipline leaders. Identifiers can encode in-repository scholars or link to common unique identifiers like ORCID and ResearcherID for dynamic updates of selection targets and scopes. Extraction and processing of literature and scholar metadata highlight scholar features to form readable, storable, associable, and displayable scholar metadata.

3.2.4 Feature Mining and Relationship Extraction

Based on selected scholars and the established logical structure framework, named entity recognition technology identifies relevant academic entities, attributes, and relationships for entity and attribute extraction. Statistical inference and reasoning based on scholar metadata mine academic attribute features, 梳理 personal identity characteristics, 计量 academic performance, and calculate relationships between scholar features to form both linear development trajectories (publication volume, h-index, work experience) and 网状 relationships (citation, collaboration, social networks). This extracts academic relationships between scholars and research entities to establish academic network models.

3.2.5 Achievement Claim and Management

After integrating scholar-related data, confirmation is required. Since libraries cannot mandate repository usage, this step targets scholars using the repository or collaborates with research management departments. Policy incentives and service upgrades guide scholars to complete achievement claims through knowledge management platforms, enabling claim, editing, modification, and statistical export of personal information and achievements. Machine learning combined with manual review validates academic achievements. For registered scholars, integrated information is pushed to their accounts for claim. Approved information undergoes feature extraction; unapproved information allows editing and reintegration into the scholar information set, creating a cyclic review and update mechanism.

4. Promotion Strategies for Digital Library-Based Scholar Repository Construction and Application

Repository construction and application constitute a complementary, iterative dynamic process following the principle of "construct while using, improve while constructing." To enhance effectiveness, strategies must motivate scholars to actively participate in information review and improvement, interface with research management platforms to improve foundational data quality, and advance applications in research process management, talent management, and resource development for management institutions, while promoting knowledge management and precision information services for scholars.

4.1 Enhance Exhibition and Publicity to Increase Scholar Participation

Due to resource quality and technical limitations, comprehensively and accurately collecting scholar information and achieving high-precision name disambiguation remain challenging. Repositories must increase scholar participation in construction and usage to improve data comprehensiveness and accuracy. Scholar pages directly relate to personal academic image, attracting scholars to enrich and maintain their information and thereby enhancing data accuracy. Employing 计量 analysis and visualization on scholar pages helps elevate

academic influence; promoting popular scholar homepages and publicizing personal display functions stimulates enthusiasm for claims and information maintenance. Access, browsing, and download behaviors generated by scholars using repository resources can serve as references for resource quality evaluation, improving repository information quality through scholar claims, applications, and interactions.

4.2 Build Research Management Platforms and Integrate into Research Processes

Integration with research management serves both management departments and repository quality improvement. Incorporating repository construction into research processes—including achievement collection, assessment, evaluation, and project application—provides foundational data for academic output submission, professional title evaluation, and research assessment, facilitating and optimizing internal research performance management and academic output statistics. Simultaneously, reviewing scholar-submitted information ensures completeness and accuracy, forming an academic information verification mechanism that strengthens academic standards and prevents research misconduct. Additionally, information in research management platforms is scholar-confirmed and highly current, serving as a data source to improve repository information quality.

4.3 Combine with Talent Identification to Serve Team and Thematic Resource Development

Feature mining and relationship extraction enable effective classification of scholars by field, discipline, specialty, or institution, identifying scholars' expertise and potential academic areas for talent identification and selection across disciplines, creating expert talent pools for review and project support. For institutional resource development, integrating specialized scholars from specific institutions or fields forms "specialized, refined, and in-depth" disciplinary thematic resource libraries that expand institutional repository distinctive resources.

4.4 Integrate with Knowledge Management to Fulfill Scholar Archiving Functions

For scholars, repositories encompassing relevant academic information and achievements serve as tools and platforms for personal knowledge management and effective storage, functioning as green repositories for open access self-archiving. They enable addition, editing, and deletion of academic achievements, transforming repositories into personal knowledge bases that not only integrate academic resources but also structurally organize scholar information to clarify academic development paths.

4.5 Construct Scholar Portraits and Models to Expand Precision Services

Embedding repositories into knowledge discovery, research management, and academic community platforms better serves scholars and research institutions. Repositories integrating multi-source scholar resources and achieving name disambiguation provide search and discovery services for scholars and their achievements. Organizing resources by scholar and 刻画 academic features from multiple perspectives provides foundational data for research management platforms, offering scholar 计量 and evaluation services. Mining scholars' disciplinary interests and development trends to construct scholar portraits and user information models approximates objective reality, laying foundations for precision resource 推送 services, recommending relevant scholars, and promoting academic exchange and collaboration.

References

[1] Zeng Jianxun. Strengthening the Construction and Application of Scholar Repositories[J]. Digital Library Forum, 2018, (9): 1-1.
[2] CNKI. Achievement Repository Help[EB/OL]. [2019-04-21]. http://papers.cnki.net/Help.aspx.
[3] Xi'an Jiaotong University Library. Institutional Knowledge Portal Open Access Policy[EB/OL]. [2019-08-27]. http://www.ir.xjtu.edu.cn/web/policies.
[4] Dou Tianfang, Zhang Chengyu, Zhang Bei. ResearcherID Status Analysis and Application Implications[J]. Library and Information Service, 2014, 58(4): 40-45.
[5] Liu Min, Tian Li. Research on University Scholar Repository Construction Path in Open Academic Environment[J]. Library Science Research, 2018(20): 28-34.
[6] TANG J, ZHANG D, YAO L. Social network extraction of academic researchers[C]//IEEE international conference on data mining (ICDM2007). Omaha, NE: IEEE, 2007: 292-301.
[7] Open academic graph|Open academic society[EB/OL]. [2019-04-15]. https://www.openacademic.ai/oag/.
[8] TANG J, ZHANG J, YAO L, et al. ArnetMiner: extraction and mining of academic social networks[C]//Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining. Las Vegas, Nevada: ACM, 2008: 990-99.
[9] Achievement Repository Help[EB/OL]. [2019-04-21]. http://papers.cnki.net/Help.aspx.
[10] Claim Achievements[EB/OL]. [2019-04-21]. http://social.wanfangdata.com.cn/index/toFistBroInfo.do.
[11] Thuid (Tsinghua University Researcherid) Service Objectives[EB/OL]. [2019-04-15]. http://rid.lib.tsinghua.edu.cn/static/web/constructionGoals.pdf.
[12] Xiao Jing, Liang Bing, Zhang Xiaodan. A Rule-Based Author Name Disambiguation Method for Document-Level Data[J]. Data Analysis and Knowledge Discovery, 2012, 28(5): 55-59.
[13] Zheng Weijie. Research on Author Name Disambiguation Methods for Scientific Literature[D]. Hangzhou: Hangzhou Dianzi University, 2017.

[14] Chang E. Research on Mechanisms and Key Technologies for Scholar Identity Identification[J]. Library Tribune, 2015(10): 88-95.
[15] Yang Wei, Lin Jing, Huang Guofan, et al. Subject Knowledge Service Marketing Strategy for "Double First-Class" Construction—Xiamen University Library Practice[J]. Journal of Academic Libraries, 2017, 35(5): 74-79.
[16] Zhang Xuelei, Wei Qingshan, Chen Yadi. Research and Practice on Scholar Academic Achievement Claim Mechanism Based on Identification Algorithms[J]. Information Studies: Theory & Application, 2018, 41(02): 68-73.
[17] Liu Wei, Zhu Zhongming, Zhang Wangqiang, et al. Research and Implementation of Author Identification and Work Claim Mechanism in Institutional Repository[J]. New Technology of Library and Information Service, 2014(3): 8-14.
[18] University of Macau Scholar Repository[EB/OL]. [2019-02-15]. http://repository.umac.mo/advanced-search.
[19] Shen Ming. AMiner: A "Magic Tool" for Scientific Research Search[N]. Science and Technology Daily, 2018-06-12.
[20] Lanzhou University Scholar Repository[EB/OL]. [2019-02-15]. http://202.201.7.4/browse-author.
[21] Scientific recruitment-hire researchers & scientists on researchgate[EB/OL]. [2019-02-15]. https://www.researchgate.net/scientific-recruitment.
[22] Li Xuan. Research on Identification Methods and Revealing Architecture for High Academic Influence Scholars in Military Medicine[D]. Beijing: Academy of Military Medical Sciences, 2017.
[23] Li Gang, Ye Guanghui. Research on Multi-Source Expert Feature Information Fusion[J]. Data Analysis and Knowledge Discovery, 2014, 30(4): 27-33.
[24] Song Peiyan, Chen Baixue, Xian Xin. Construction and Empirical Research on Semantic Model for Scientific and Technical Expert Information[J]. Information Studies: Theory & Application, 2017, 40(9): 119-124.
[25] Fan Xiaoyu, Dou Yongxiang, Zhao Pengwei, et al. Research on Construction Method of Researcher Portrait Integrating Multi-Source Data[J]. Library and Information Service, 2018, 62(15): 31-40.

Author Contributions:
Zheng Ang: Responsible for topic investigation, analysis, and paper writing;
Zeng Jianxun: Proposed the topic and ideas, revised the paper.

Study on the Construction Method of Scholars Repository Based on Digital Library

Institute of Scientific and Technical Information of China, Beijing 100038
Zheng Ang, Zeng Jianxun

Note: Figure translations are in progress. See original paper for figures.