

Intelligent Information Systems and Digital Libraries

Yuqing Shi^{1, 2, a}, Yuelong Zhu^{2, b}

¹ Library, Hohai University Nanjing, China

² College of Computer and Information Engineering, Hohai University Nanjing, China

^ashiyuqing@hhu.edu.cn, ^bylzhu@hhu.edu.cn

Keywords: Digital libraries; Tailored Services; Intelligent Information Systems; Data Detection

Abstract. Digital library information systems generate large amounts of complex, numeric content. A new category of digital libraries with supporting services for such content is important to advancing library research. Intelligent services likewise are required to automate scientific workflows. Simulation environments are notably well suited for automation with digital libraries. This paper shows a new method to build intelligent information systems framework generates simulation supporting digital library instances. This framework contains processes for producing workflows of layers and services, formal definitions of content and services, a software toolkit, and a set of service implementations. Coordinating such digital library examples within a library practice infrastructure leads to more intelligent services and to the automation of numerous routine tasks. The intelligent information systems are is developed at library of Hohai University in China.

Introduction

A new generation of intelligent information systems will appear from digital library study. One side of this is based on intelligent applications and technologies in the digital library field. Another side is based on progress in digital library theory and related advances upon advanced integration and quality. Together, these two sides promote automatic operation and empower users. Since the incipency of the domain of digital libraries in 1990s, there has been a stress on advancing the state of the art in intelligent information systems; in the immediate future that also should aim to more tailored and smarter services [1]. Intelligent information systems may be looked as filling the gap between knowledge management systems and database management systems [2]. They most generally have been implemented to multimedia content and text, but also make use of more strictly structured data, as same as graphs and a great variety of representations that support inference and carry semantics [3].

Digital libraries can become wiser through some types of connections with advanced technologies. One method is for received content to be analyzed wisely, as through the usage of natural language treating, heuristic based mining, detection, extraction and parsing, social network structure, conversion into metadata generation, and canonical forms. Another method is for the appease they work with to be received wisely, as through focused crawling. Value can be added to the set of representations in a digital library as much through detecting, annotating, classifying, clustering, and summarizing. A second set of technologies meets with tailoring and enhancing the structure, archival, and maintenance operation of digital libraries [4].

Digital libraries had led to continuations into, and upgrades leading to, new types of information systems [5]. One sample is repositories building on systems like tailored Hydra solutions. Another sample is courseware management systems, with open source solutions such as Sakai and its offshoots [6]. A third sample is the sparkly field of content management systems, with powerful open source environments [7]. A fourth sample is personal information management systems. These samples make fine that digital libraries advances toward more intelligent systems can lead the way to a wider area of systems, assorting from the personal to the organization as same as the global, and from the small to the large.

The rest of the paper is organized as follows: the next segment introduces the digital libraries for science research and third segment presents the science supporting digital libraries. Fourth segment describes the intelligent information system production framework, and finally conclusions are consisted in the last segment.

Digital Libraries for Science Research

Nowadays, scientific study practices increasingly desire data-intensive information management and high performance computing. Computational science workflows generate a good deal of data. Undertaking back this type of study, there are tiny best tools or practices for individual researchers and organizations to accomplish for orderly taking, using, and managing data products. Mixing this conundrum, it is hard to produce normal metadata schemas for scientific areas to let interoperable between efforts and study infrastructures. Study teams usually lack the expertise, incentive, and drive to appropriately catalog produced content and model. The lack of validity of tools and semantic system has led to study infrastructures that lack intelligent services for searching, sharing, and collaborating.

In science research, for example, emulation software packages and modeling, there always is a necessity for calculating resources including big data and high performance computing systems. These packages generate petabyte scale researches of scientific data that may be lessened by brief to several terabytes of aggregated data. There are few effective digital library software applications for automating packages science supporting tasks, enabling secondary purpose study with content, for example, data detection, multi institution integration, data mining, and reusing, or focusing investigators' time on high level tasks. Thus, the information management systems ask advanced services for preserving and archiving content.

Several simulation and modeling applications motivated trials to produce a new category of simulation supporting digital libraries that afford higher level services. Nowadays, simulation and modeling is used in many areas to research and represent the dynamics of real world scenarios. Some simulation applications may be defined by the output results generated, methodic input parameter collocations necessary, as same as defining schemas for the known. The simulation based study workflows performed by a study system may be regularly described by defining the stages of content and processes that map or transform between continuing content stages. As a sample, a methodic simulation based workflow might be formed from content stages including input configurations, annotations, visualizations, figures and plots, results, statistical analyses, summarized results, publications, and application schemas. Simulation based scientific study be formed from processes that could be semi automatically enabled or conducted via well managed scientific collections, supported by digital libraries.

Digital librarians have a long time of offering information retrieval and storage services in multiple full text areas. Nevertheless, these higher level services desire new metrics and algorithms to exactly support big data and highly numerical scientific content. The precondition of simulation supporting and rising scientific digital librarians is that made-to-order services enable users to operate at a higher level to the degree that intelligent behavior is structured into those higher level services.

Science Supporting Digital Libraries

In this paper intelligent services were defined as software components that afford intelligent digital librarian functionality, interoperable with other services, and preparation into workflows of higher order abstract services. The preparation stress of intelligent services permits for low level as same as complex services to be used as part of other services. Digital librarians have the feasibility to help establish scientific collections and restructure automated data management tasks for formerly uncorroborated scientific areas. Push-button workflows of intelligent services lead to and support a focusing of research fellow trials on higher level tasks. With suitable joints to high performance

resources and software, digital librarians can automate the achievement of scientific, structured, pre-determined workflows.

Formal definitions of digital librarians are made up to mathematics based, precise instructions of societies, services, content, and along with their correlations. Societies are defined by computing actors and grouping human based upon services, responsibilities, goals, tasks, scenarios, and concerned content. Service definitions are based on their management signature of formally defined output and input content structures. Content is defined based upon representations of content structures and byte streams. A formal model for digital libraries is named 5S that included in: streams, structures, spaces, scenarios, societies. 5S means digital libraries were defined as complex information systems that communicate information with users, organize information in usable ways, present information in usable ways, provide information services, and help satisfy information needs of users. Combinations of formal definitions from each of the 5S areas code the varied aspects of a digital library. Digital librarians were viewed from the 5S viewpoint that digital librarians can lead to many future advances:

Formal definitions of digital library services are gone with added informal descriptions to consider human understanding and mathematical reasoning. For each service, a formal definition includes the required pre-conditions, implementation details, post-conditions, inputs, and outputs. In this paper, usage of the 5S formal framework supplies a technology to formalize the implementation details, structure, and design for broadly needed services.

The following services are necessary for a lowest simulation supporting digital library. A low-level metadata extraction service generates a set of digital objects that may be used by a metadata register production service. Curation is an unclear, complex process made up of deciding which content to sustain and under what duration and conditions to keep it. By means of new automated common curation services, the curation policies within a deployed system may still be made-to-order to a specific field of science. The curation services is supplied a set of digital objects and components including curation operators, grammars, rules, catalogs, metadata records, handles, and archived documents. Via origin backtracking, higher level provenance tracking services may detect the generation process and genetic source for a given digital object. The provenance tracking service maintains and builds a metadata index of boundaries describing the origin details between pairs of participant digital objects to get the stages of and relationships between content.

The definitions for the lowest scientific digital library afford a formal specification of the claims needed in this domain. The definitions were used to generate a software toolkit for each of the lowest required services including metadata extraction services, curation services, and provenance tracking services. This software toolkit was then utilized to achieve cases of the digital library software modules for modules instance researches.

Intelligent Information System Production Framework

The systematic outcome of digital libraries with automated, tailored functionality needs a structured production framework. The intelligent information system production framework uses a digital library software toolkit to generate tailored cases with science supporting, automated functions. The automated services are included in metadata extraction, digital objects, metadata record production, metadata record indexing, ranking, filtering, search, communication, staging, contribution tracking, incentivizing, and so on. The 5S formal framework is used to generate formal definitions of a digital library system's marginally necessary societies, scenarios, spaces, structures, and streams. Formal definitions are used by digital library artists as a basis for encoding and formulating scientific workflows, types, content stages, and services.

In this paper, a complete digital library is developed by stepping through the layers of the model including digital library management system, digital library system, and digital library instance. The representation of the intelligent information system production framework is include in formalisms in a digital library management system, software toolkit in a digital library system, and produced digital library instances. The tasks and processes supported including manual metadata term-value

production, value, entering, and completeness checking, identifying duplication; sharing and reusing content, copying, loading, and transferring content, moving data from one place to another; initiating successive services with intermediate data objects providing statistics regarding software, models, and content usage, providing customized reports to contributors of digital objects, software, and high-performance computing systems and so on. Digital library functionality is gotten from the usage of definitions and the software toolkit consisting of field specific and routine services. This field produced intelligent system level behavior, consisting of high level services that automate ago human concentrated tasks.

Summary

Nowadays, there has been great development over more than 20 years of work inside the digital library domain to produce more intelligent systems. By using a formal method, and addressing pivotal problems of quality and comprehensiveness, methods based on intelligent approaches can be used in a broad range of application areas. This paper shows a new method to build intelligent information systems that is developed at library of Hohai University in China.

Intelligent digital libraries will produce enhancements to the embranchment from digital library work, such as digital library information systems, and digital library content management systems, as same as digital library institutional repositories. And then, the digital library will field higher level human machine synergy effect, and will enhance software engineering practices utilized in this field. To human users who favor human computer intergrowth, intelligent support will be supplied to those who digital libraries, operate, build, and design future digital libraries.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (No. 51079040/E090101) and Hohai university library '2015' key research projects (No.20150323).

References

- [1] Zhao Yumin and so on. Research on Data Mining Technologies for Complicated Attributes Relationship in Digital Library Collections [J]. Applied Mathematics & Information Sciences .2014:8 (3):1173-1178.
- [2] Li Yao and so on. Shared and Service-oriented CNC Machining System for Intelligent Manufacturing Process [J]. Chinese Journal of Mechanical Engineering .2015:28 (6): 1100-1108.
- [3] Gross Geoff A. and so on. Precedence tree guided search for the efficient identification of multiple situations of interest - AND/OR graph matching [J]. Information Fusion .2015:27: 240-254.
- [4] Shoaib Muhammad and so on. Improving Similarity Measures for Publications with Special Focus on Author Name Disambiguation [J]. Arabian Journal for Science and Engineering .2015:40 (6): 1591-1605.
- [5] Javier Cabrerizo and so on. A decision support system to develop a quality management in academic digital libraries [J]. Information Sciences .2015:323: 48-58.
- [6] Yu Wei and so on. Automatic Evaluation for Engineering Pathway Premier Award Winners [J]. Applied Mathematics & Information Sciences .2014:8 (5): 2613-2626.
- [7] Kogalovsky M. R. and so on. Social network technologies for semantic linking of information objects in scientific digital library [J]. Programming and Computer Software .2014:40 (6): 314-322.