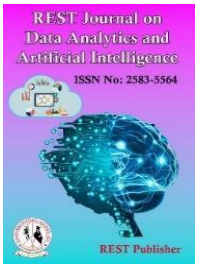
REST Journal on Data Analytics and Artificial Intelligence
Vol: 2(3), September 2023
REST Publisher; ISSN: 2583-5564
Website: http: //restpublisher.com/journals/jdaai/
DOI: https://doi.org/10.46632/jdaai/2/3/6



A Survey of Bigdata Analysis, Extracting Data and Mapping the Data

[*1]P. Hemalatha, [2]J. Lavanya
[1]*Adaikalamatha college of Arts & Science College, Thanjavur, Tamil Nadu, India.*
[2]*Dr. Nallikuppusamy Arts & Science College, Thanjavur, Tamil Nadu, India.*
[*]Corresponding Author Email: saikirthik87@gmail.com

Abstract: *Data mapping is one of the simplest terms is to* map *source* data *fields and their related target* data *fields. Mapping can have a varying degree of complexity, depending on the number, data types, schema, primary keys, and foreign keys of the data sources. Nowadays, Archaeological research is based on an interdisciplinary approach which makes use of a wide range of technologies allowing for the collection of data and information about sites and archaeological findings. The purpose of archaeology is to learn more about past societies and the development of the human race. An essential part of the archaeological data is related to spatial information that links historical contents to the metric reconstruction of monuments and artifacts, and show their mutual relations in a map. A critical a part of the archaeological records is associated with spatial data that links ancient contents to the metric reconstruction. By processing a steady stream of all real-time data, organizations can make time-sensitive decisions faster than ever before, monitor emerging trends, course-correct rapidly and jump on new business opportunities. To design a data mapping framework process, the data from various sources uses a new proposed technique. To secure the high profile raw and analyzed data using the combination of hardware and software any key generation for data extraction and mapping. The information can be accessed only through the authenticated source of the framework and hence duplication and data theft is extremely difficult. This paper follows the various data mapping techniques handled in previous work and also shows the limitations of existing techniques.*
Keywords: *Data mapping, cleaning techniques, data extraction, Bigdata*

1. INTRODUCTION

Big data refers to data sets that are too large or complex to be dealt with by traditional data-processing application software. Data with many fields (rows) offer greater statistical power, while data with higher complexity (more attributes or columns) may lead to a higher rate. Big data analysis challenges include capturing data, data storage, data analysis, search, sharing, transfer, visualization, querying, updating, information privacy, and data source. Big data was originally associated with three key concepts: volume, variety, and velocity. The analysis of big data presents challenges in sampling, and thus previously allowing for only observations and sampling. Therefore, big data often includes data with sizes that exceed the capacity of traditional software to process within an acceptable time and value. Big data analytics helps organizations, belonging to any sector, harness their hard-earned data and employ it to identify new opportunities that are hidden previously. That, in turn, leads to better business investments, more accomplished operations and higher remuneration. The following figure (Figure 1) gives a clear illustration of the usage of Big data Analytics.

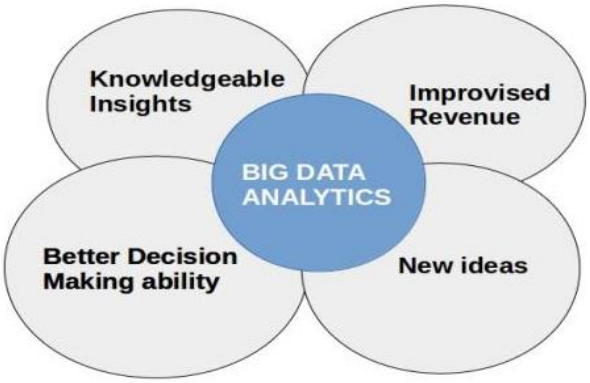


FIGURE 1. Usage of Big data Analytic

2. DATA MAPPING

Data mapping is an essential part of many data management processes. If not properly mapped, data may become corrupted as it moves to its destination. Quality in data mapping is key in getting the most out of your data in data migrations, integrations, transformations, and in populating a data warehouse. Data migration is the process of moving data from one system to another as a one-time event. Generally, this is data that doesn't change over time. After the migration, the destination is the new source of migrated data, and the original source is retired. Data mapping supports the migration process by mapping source fields to destination fields.

Types of data mapping

2.1. Manual Data Mapping

Manual records mapping calls for expert coders and records mappers — IT will code and map your records datasets. Although that is a heavy carry and calls for expert help, it lets in you to completely manage and personalize your maps.

2.2. Semi-automatic Data Mapping

Semi-automatic records mapping (or schema mapping) calls for a few coding understanding and way your group might be transferring among each guide and automatic records mapping processes (as a result the call of this technique). Data mapping software program creates a connection among the records datasets after which an IT expert opinion the ones connections and makes guide changes as needed.

2.3. Automated Data Mapping

Automated records mapping way a device will cope with all components of the records mapping method for you, making it a perfect option, in case you are not/ don't have get admission to a coder. This sort of software program will commonly permit for drag-and-drop mapping. You simply want to discover ways to use the device (and pay for it). Speaking of the gear so one can automate the method of records mapping for you, let's evaluation a number of your alternatives next.

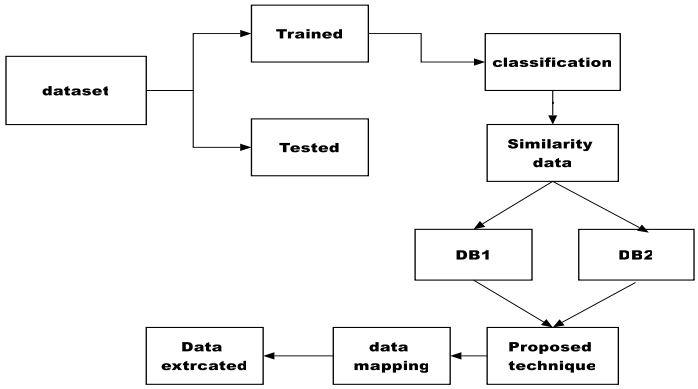


FIGURE 2. Flow diagram of Data Mapping

The data mapping and pruning (cleaning techniques) are crucial for effective use of big data because they can find the hidden relation among items. Such relation can help us find unwanted or useless data's which can be modified accurately. The Proposed technique can be considered as special forms of quality rules are the core of effective data cleaning systems with apply on large and low-quality datasets. To handle the volume issue of big data, this research idea develops the sampling algorithms to obtain a small representative training set. It formulates the design the fault-tolerant rule discovery and conflict-resolution algorithms to address the low-quality issue of big data. It also proposes parameter selection strategy to ensure the effectiveness of CFD discovery algorithms. Conditional Functional Dependency (CFD) discovery algorithms are powerful tools in data cleaning, because they can find the hidden relation among items. Such relation can help us find dirty tuples which can be modified accordingly. The Functional Dependencies (FDs) can be considered as special forms of CFDs. High-quality rules are the core of effective data cleaning systems with CFDs. Current Conditional Functional Dependency (CFD) discovery algorithms always need a well-prepared training dataset. This condition makes them difficult to apply on large and low-quality datasets. To handle the volume issue of big data, this research idea develops the sampling algorithms to obtain a small representative training set. It formulates the design the fault-tolerant rule discovery and conflict-resolution algorithms to address the low-quality issue of big data. It also proposes parameter selection strategy to ensure the effectiveness of CFD discovery algorithms. Experimental results demonstrate that our method can discover effective CFD rules on billion-tuple data within a reasonable period.

3. ARCHAELOGY

In archaeology, survey or field survey is a sort of field research by which archaeologists (frequently scene archaeologists) look for archaeological locales and gather data about the area, circulation and association of past human societies across an enormous region (for example commonly more than one hectare, and frequently in abundance of numerous km2). Archaeologists lead surveys to look for specific archaeological locales or sorts of destinations, to identify designs in the conveyance of material culture over areas, to make speculations or test theories about past societies, and to evaluate the dangers that improvement undertakings will unfavourably affect archaeological heritage. The surveys might be: (a) nosy or non-meddlesome, contingent upon the requirements of the survey group (and the gamble of obliterating archaeological proof assuming nosy strategies are utilized) and; (b) broad or serious, contingent upon the kinds of exploration questions being asked of the scene in inquiry. Surveys can be a viable method for choosing whether or not to do an uncovering (as an approach to recording the essential subtleties of a potential site), yet may likewise be closes in themselves, as they produce significant data about past human exercises in a local setting. A typical job of a field survey is in evaluation of the possible archaeological meaning of where improvement is proposed. This is normally associated with development work and street building. The evaluation decides if the area of advancement sway is probably going to contain critical archaeological assets and makes suggestions concerning whether the archaeological remaining parts can be kept away from or an unearthing is fundamental before improvement work can begin. Archaeologists utilize an assortment of apparatuses while completing surveys, including GIS, GPS, remote detecting, geophysical survey and aeronautical photography.

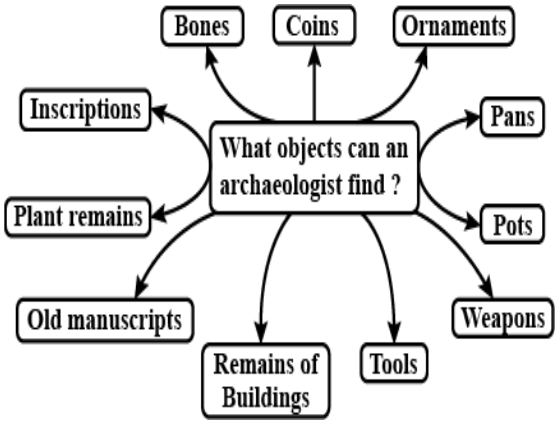


FIGURE 3. Various types of archaeologist data

4. RELATED WORK

Andrea Gennaro; Michele Mangiameli[1] This paper describes a variety of modern geomatics techniques for the detection and survey of emerging archeological structures in Sicily. The case study presented in this work is related to an archaeological site that lies on the north-western flank of Etna, the highest active volcano of Europe. For the field survey, we used an Unmanned Aerial Vehicle (UAV) equipped with a high-resolution camera and a GPS for the ground control points. In addition, we produced a Digital Surface Model (DSM) and several detailed orthophotos of the detected objects. The geomatic approach here adopted made significant contributions to our reconstruction of the past settlement located between "Balze Soprane" and "Santa Venera" districts. Xiaoqian Chen School of Marxism [2] Ancient bronze lifting cylinder is one of the important bronzes of the Bronze Age culture in Vietnam and parts of southern China. It is special in shape, beautifully decoration and rich in content. Based on the data of Archaeological institutes, museums and related cultural institutions, this paper mainly used archaeological methods to analyze the types, decoration and functions of ancient bronze lifting cylinder. At the same time, it combined historical and ethnological materials to conduct comprehensive research. The conclusion is as follows: bronze lifting cylinder is one of the typical utensils in the Bronze Age Culture in Vietnam and parts of southern China. It is the most representative of the Bronze Age culture in Vietnam. It may originate from the Dongtan culture in Vietnam and Prevailing in here. Through crowd migration or diplomacy, a small amount spread to parts of southern China, and on the basis of the addition of local elements, it became a unique artifact in its bronze culture. It not only shows the relatively mature Bronze Age culture in this region, but also provides us with relatively straightforward materials for studying its cultural connotations Maria Costa; José Pinto; Manuel Ribeiro[3] .This specific vehicle includes sonars, an optical camera and a magnetometer to detect and identify archaeological artifacts in the ocean bottom or underneath. It can be used isolated or as part of a team of AUVs for faster surveys. We describe the hardware, its simplified operation using custom-made software and overview some results in different areas, where these vehicles have been used to detect several wrecks and other important artifacts The main innovations used in our work for simplified archaeological surveys in coastal areas are man-portable and re-configurable AUVs, as well as simplified operations through in-house developed software, together with self-compensating magnetometers from Ocean Floor Geophysics. These tools and technologies are a good starting point towards answering the question we propose to address in this work. Alejandro Alvarez-Ayllon [4] This paper follows the guidelines for systematic mapping studies, which is well suited for gathering and classifying available studies. We summarize the results after classifying the 242 papers that passed our inclusion criteria. While there are many proposed solutions that tackle the problem in different manners, there is little evidence available about their implementation in practice. Almost all of the solutions found by this paper cover a subset of our requirements, with only one partially satisfying the three. The solutions for data exploration abound. It is an active research area and, considering the continuous growth of data volume and variety, is only to become harder. There is a niche for research on a solution that covers our requirements, and the required building blocks are there. G. Vallicrosa, M. J. Fumas, F. Huber [5] The present work tries to improve the topographical and prospective methodology applied to underwater archaeology, as well as the different technical tools used in the study of submerged caves, using as a experimental site one of the most contaminated and dangerous submerged caves in Europe called Falconera Cave (Sitges, Barcelona). The results have shown that the application of traditional procedures together with the use of the Sparus II Autonomous Underwater Vehicle (AUV) equipped with a multibeam echosounder represent the best technical solution for georeferencing in these submerged archaeological contexts David Novák; Martin Kuna; Olga Leèbychova [6] The AIS CR infrastructure consists of a central database and several linked application interfaces. It enables the day-to-day management of archaeological data and fieldwork activities, supports research and development, and promotes the nation's archaeological heritage. Equipped with the latest technologies and open-source software, the AIS CR provides both professionals and the wider public with open access to a vast array of high-quality and linked archaeological data.

5. CONCLUSION

Big data investigation can possibly change the way which demonstrates modern advances to pick up knowledge from their clinical and other information stores and settle on educated choices. There are numerous instruments

accessible to investigate the current day enormous information in a secured way however our examination has demonstrated that the proposed coordinated explanatory system with installed security is more proficient with altered .data record arrange which helps the examination procedure and the security is saved with the two-level verification of programming and equipment confirmation. The proposed work mainly discussed about the data extraction in big data by analytics for archeological data. Mapping and data extraction with time-consuming and non-trivial exercises with great potential for error. This research follows the data mapping with anchelogical data and overcome all the limitations.

REFERENCES

[1]. Andrea Gennaro; Michele Mangiameli et.al"Geomatic techniques for surveying and mapping" an archaeological site Published in**:** 2019 Metrology for Archaeology and Cultural Heritage (MetroArchaeo)IEEE Xplore: 11 May 2020.

[2]. Xiaoqian Chen School of Marxism, Yunnan Normal University, Kunming, China "Archaeological Research of Ancient Bronze Lifting Cylinder" 2019 3rd International Conference on Data Science and Business Analytics (ICDSBA) IEEE Xplore: 30 November 2020

[3]. Maria Costa; José Pinto; Manuel Ribeiro et.al" Underwater Archaeology with Light AUVs" Published in: OCEANS 2019 - Marseille IEEE Xplore: 14 October 2019.

[4]. Alejandro Alvarez-Ayllon, Manuel Palomo-Duarte2, And Juan-Manuel Dodero (2019) sInteractive Data Exploration of Distributed Rawfiles: A Systematic Mapping Study -Ieee Access Volume 7.

[5]. G. Vallicrosa, M. J. Fumas, F. Huber and P. Ridao, et.al "Sparus II AUV as a Sensor Suite for Underwater Archaeology: Falconera Cave Experiments," 2020 IEEE/OES Autonomous Underwater Vehicles Symposium (AUV), St. Johns, NL, Canada, 2020, pp. 1-5, doi: 10.1109/AUV50043.2020.9267935

[6]. David Novák; Martin Kuna; Olga Leèbychova"et.al IT and the Humanities in the 21st Century — The Case for Archaeology " Published in: 2019 9th International Conference on Advanced Computer Information Technologies (ACIT) Date Added to IEEE Xplore: 01 August 2019.

[7]. Mohit Mittal;Panote Siriaraya;et.al "Accurate Spatial Mapping of Social Media Data with Physical Locations" Published in**:** 2019 IEE International conference on bigdata IEE Xplore:1 May 2020.

[8]. A. Sutton, M. Clowes, L. Preston and A. Booth, "Meeting the review family: Exploring review types and associated information retrieval requirements", Health Inf. Libraries J., vol. 36, no. 3, pp. 202-222, Sep. 2019.

[9]. H. Liu, M. Zhou and Q. Liu, "An embedded feature selection method for imbalanced data classification", IEEE/CAA J. Automatica Sinica, vol. 6, no. 3, pp. 703-715, May 2019.

[10]. A. Chapman, E. Simperl, L. Koesten, G. Konstantinidis, L. D. Ibáñez, E. Kacprzak, et al., "Dataset search: A survey", VLDB J., vol. 29, no. 1, pp. 251-272, 2020

