Nat. Volatiles & Essent. Oils, 2021; 8(4): 11811-11834

NVEO
Natural Volatiles & Essential Oils

Measurement Model In Optimizing The Certainty Of Accurate Data Similarity With The Hierarchy Of Partition Grid Method (HGP)

Verdi Yasin[1*], Muhammad Zarlis[2], Opim Salim Sitompul[3], Poltak Sihombing[4]

[1]Graduate Doctorate Program in Computer Science, Universitas Sumatera Utara, Medan, Indonesia

[2]Department of Computer Science, Universitas Sumatera Utara, Medan, Indonesia

[3]Department of Information Technology, Universitas Sumatera Utara, Medan, Indonesia

[4]Department of Computer Science, Universitas Sumatera Utara, Medan, Indonesia Jalan Universitas Nomor 9, Padang Bulan, Medan Baru, Medan, Sumatera Utara 20222 Indonesia

ORCID: https://orcid.org/0000-0002-1967-5046

Abstract:

The development of the increasing population, the number of student academic data population is currently increasing, making researchers interested in conducting research in order to ensure student entities, especially validation of the certainty of student data, including ways to facilitate the process of finding student data, when verifying data certainty such as related attributes in a student, among others:Student Parent Number. Student Name, Place of Birth, Date of Birth, Gender, Study Program, College Name, Diploma Number, and other supporting data. The measurement analysis model used in this research method is the Hierarchy of Partition Grid (HGP) data integrated in the database system, in order to facilitate verification and validation through smart phone numbers or through the internet network, using the method of hierarchy of partition grid (HGP).As a solution to the above problems, the researcher formulated a modeling of an integrated data processing system between the attributes of academic data of students, so that it is easy to get information related to academic entities of student records. In describing the concept of this integrated data processing model. Researchers also use object-oriented analysis with unified modeling language (UML) design tools, in addition to using structured data relationship modeling entity relationship diagram (ERD).From the above explanation it can be concluded that the processing of student academic data, it is necessary to create a proper model to overcome the difficulties in ensuring student academic data at a college.The results of measuring the accuracy of data similarity using the Hierarchy of Grid Partition

(HGP) population data for student academic data, from a total of 5185 data, it can be seen that the measurement results that amounted to 4111 normal data (single ID), amounted to 1070 invalid double data (similarities), and amounted to 4 valid double data (similarities), so it can be concluded that there are 4 data that have a valid level of accuracy of similarity.

INTRODUCTION

Along with the development of the number of residents in the Republic of Indonesia, student data in a college or university is increasingly increasing data every year, so the academic system will experience a problem when ensuring attributes to student data entities in the search process because there is a similarity of data between universities. Therefore, it is necessary to integrate data from population data, along with the increasing ability of database technology, it is necessary to measure the right data attribute record to test the correctness of attributes contained in the academic database system, especially attributes that are related to the academic record system and related to other academic data records such as primary data, secondary data and student academic support data [1]-[4], [30]-[31].

From several interrelated sections about students' academic records must be integrated into the database system. The very basic part in the research is that there are many difficulties in accumulating student academic data, Both those who have not graduated or who have graduated, it is necessary to create a proper and accurate model in the integration of student academic data, including supporting data about students' personal, must be integrated into one system in order to be controlled easily and effectively[5]-[13].

Measurement Model with Hierarchy of Grid Partition (HGP) academic data that can find student data that has similarities or double data, in addition will also be able to menu bad data or students who do not have certainty of data truth [30]-[31].

The purpose of the Measurement Model with Hierarchy of Grid Partition (HGP) academic data is to optimize the processing of student data in a college, by integrating student population data with academic data, so that when searching for student data, it can use population data with Student Population Identity Number [30]-[31].

The processing of academic data is integrated in a database, in order to facilitate the search for academic data of students, thus the potential for falsification of student academic data can reduce, even reduce the occurrence of the creation of fake Academic Graduate Diploma data by certain individuals.

MATERIALS AND METHODS

Material Review

Measurement Models with Hierarchy of Grid Partition (HGP) academic data can be described using a variety of notations, such as unified modeling language (UML). In UML notation, unified modeling language diagrams can be used in these measurements as class diagrams, where classes represent data attributes in a relationship between student data attributes, An association represents the relationship between attributes and the type of role of an association representing the type of role taken by examples of measurement models in various situations. In Entity Relationship (ER) notation, the conceptual model is described by the Entity Raltionship Diagram (ER) in which the entity represents the data object, cardinality and optionality represent the relationship between the attributes of the object (data entity) [1]-13,[10]-[21].

System development is a way to update the system to be new in replacing the old system as a whole or improving the existing system. System development methodology is a standardized process followed by organizations to implement all the steps necessary to analyze, design, implement and maintain information systems. There are various methodologies of system development, alternative solutions by choosing the best solution. Then create a process design for the data that has been obtained and modeled in the database system architecture. The tools used in system modeling usually use UML (Unified Modeling Language) [16]-[20].

Data processing is the manipulation of data in order to become a more useful form. This data processing is not only a numerical calculation but also operations such as data classification and data transfer from one place to another. In general, we assume that these operations are carried out by some type of machine or computer, although some of them can also be done manually [1]-[10].

Data processing consists of three main steps, namely input, process (processing), and output [1]-[21].

- **Input:** In this step the initial data, or input data, is prepared in some form suitable for processing purposes. The form will depend on the processing of the machine.
- **Process:** In this step the input data is changed, and is usually combined with other information to produce the data in a more usable form. This processing step usually includes a series of specific basic processing operations.

- **Output:** In this step the results of the previous processing are collected. The form of output data depends on the use of the data for further processing.
- **Origination:** This step is the process of collecting original data (original / raw data). The original record of this data is called the source of the document. For example, the source of the document from the example of the niali calculation case above is the test file of students who have been assigned a grade. Please note that if there is a statement regarding the final grade of the student, we can look back at the source document (the student test file) and check for possible errors that have been made during this step.
- **Storage**: This step is a very important step in any data processing procedure. Data processing is often placed in storage for use as input data to be processed at the next time. Two arrows in the process box and storage box show the interaction of these two steps. A set of data that forms a single entity in storage is called a file. Usually a file consists of a set of records, where each record contains the same data item. Next is a collection of interconnected files.

Data processing procedures usually consist of a number of basic processing operations carried out in several sequences [20]-[29].

- **Recording**: Recording is moving data on multiple forms or documents. This occurs not only during the origination stage (in the source document) and the distribution stage (on the report document) but occurs throughout the processing cycle.
- **Duplicating:** This operation is a doubling of the data on top of forms or documents. Duplication may be done when the data is recorded manually, or it may be that duplication is done afterwards using a machine.
- **Verifying:** Since recording is usually a manual operation, it is important that the data that has been recorded is carefully examined, there may be errors.
- **Classification :** This operation separates data into different categories. Classification can usually be done in more ways than one. For example, a set of student question lists can be classified according to the gender of the student, or according to the student's year of admission.
- **Sorting:** Organize data in a specific order. This operation often occurs in everyday life. The names in the phone book are sorted alphabetically, employee data is disorting according to the employee's parent number. Sorting data can be done before or after classification.

Example: a Student file contains data items: Student Parent Number, Population Master Number, Name, Gender, Religion, study program, year of entry, year of exit, Diploma number, college name. Sorting can also use more than one sorting key, i.e. with the first key, the second key and so on. First sort based on the first key when there is data similarity then the second key is used and so on [1]-[10], [20]-[29].

- **Merging:** This operation is to mix two or more data sets, all of which have been linked with the same key, and put those data sets together into a single data set that has been disorted.
- **Calculations:** Perform numerical calculations on numerical type data.
- **Check table:** search and regain data (table look-up, searching, retrieing). This operation intends to recover certain data in the data set that has been investigated.

Research objectives

This research aims to create a formulation of data integration models in the database to facilitate the process of accessing or searching for student academic data at a college, when proving the truth of student data and allowing accurate student data. In addition to avoiding the occurrence of the misuse of diplomas by certain individuals, when a person commits an act of forgery against the Diploma with the same name and the same parent number, but other identities are different. This integrated data processing concept, all attributes related to the student will be indexed by integrating all attributes about the student through primary key,forein key and alternative key [1]-[21].

Research Methods (Measurement Model with Hierarchy of Partition Grids (HGP)

To complete this integrated measurement of academic data, it needs an appropriate method, so that all processes can be integrated properly and effectively. So the researchers used the Hierarchical Method of Partition Grid (HGP) [30]-[31].

Researchers took the data set as a sample of data in this measurement amounting to 5185, is attached as follows:

Table 1. Table data set, as sample data in the measurement of this academic data:

ID	POPULATION IDENTITY NUMBER	name	PLACE OF BIRTH	DATE OF BIRTH	STATUS	GENDER	ADDRESS	RT	RW
513313	3171035408880003	DWI PUTRI ADRULINA	JAKARTA	14|08|1988	S	P	JLN RANJAU NO 9	001	005

513314	3171030204980002	MUH.ALFIAN FADHILA	JAKARTA	02|04|1998	B	L	JL. PUALAM V NO. 173	017	002
513315	3171035205990006	MELIANA	JAKARTA	12|05|1999	B	P	JL BERLIAN RAYA NO.179 A	017	002
513316	3171034806010004	RIA YUNIARTI	JAKARTA	08|06|2001	B	P	JL PUALAM 5 NO 176	017	002
513317	3171024201010003	RORO ANDI GANEESA PUTRI	JAKARTA	02|01|2001	B	P	SUMUR BATU	017	002
513318	3306025608760002	NURMI INTAN SUDIHATI	PURWOREJO	16|08|1976	S	P	JL. PUALAM RAYA NO 21	017	002
513319	3201135309890009	ANDI SYARIFA PRATIWI	UJUNG PANDANG	13|09|1989	B	P	JL. PUALAM V NO 172	017	002
513320	3172050210931001	YUDI PURNAMA	JAKARTA	02|10|1993	S	L	JL. SUMUR BATU	016	002
513321	3171036607930003	YULIANA	JAKARTA	26|07|1993	S	P	JL. SUMUR BATU	016	002
513322	3327015912920005	IRMAWATI	PEMALANG	19|12|1992	S	P	JL. SUMUR BATU	016	002
513323	3276026704860012	YESSY	JAKARTA	27|04|1986	S	P	JL. SUMUR BATU NO 41 A	016	002
513324	3275126001900007	ESTERINA NURIDAYANTI DEVI	JAKARTA	29|01|1990	S	P	JL. PUALAM RAYA NO 35	016	002
513325	3275055510740022	EVI DARWATI	GUNUNG	15|10|1974	S	P	JL.	016	002

			KIDUL				PUALAM II NO 87		
513326	3171034812870004	SHAILLA SITUMORANG	JAKARTA	08|12|1987	S	P	JL. PUALAM I NO 48	016	002
513327	7371125301900007	ERWINDA PERDANANINGSIH	MAKALE	13|01|1989	S	P	JL. PUALAM II NO 112 D	016	002
513328	3528042306880003	MOHAMMAD NABIL	PAMEKASAN	23|06|1988	S	L	JL. SUMUR BATU RAYA	016	002
513329	3174030401840007	DWI WINARTO	JAKARTA	04|01|1984	S	L	JL. SUMUR BATU RAYA	016	002
513330	3175096409840002	SUHARTINI	RANGKAS BITUNG	24|09|1984	S	P	JL. SUMUR BATU RAYA	016	002
513331	3175022508910002	ACHMAD JAYA WINANGUN	JAKARTA	25|08|1991	S	L	JL. SUMUR BATU RAYA	003	001
513332	3275032309870007	YOGIE ARDHANA	DUMAI	23|09|1987	S	L	JL. SUMUR BATU RAYA	003	001
513333	3171020202820005	FERYADI	JAKARTA	02|02|1982	S	L	JL. SUMUR BATU RAYA	003	001
513334	3174040210850001	ROBIT HARYADI	JAKARTA	02|10|1985	S	L	JL. SUMUR BATU RAYA	003	001
513335	3171036805930001	MARTI ENDA SAPITRI	JAKARTA	28|05|1993	S	P	JL. SUMUR BATU RAYA	003	001
513336	3171035005920003	SRI SUTARI	JAKARTA	10|05|1992	S	P	JL. SUMUR BATU RAYA	003	001
513337	3307064507850001	PAINAH	WONOSOBO	05|07|1985	S	P	JL. SUMUR BATU RAYA	002	001
513338	3171030910001001	MUHAMMAD FARHAN ARFPIA	JAKARTA	09|10|2000	B	L	JLN MORTIR II / 3 KODAM	004	005

							SUMUR BATU		
513339	3171036409010002	KHANSA MAHIRA	JAKARTA	24|09|2001	B	P	JL. GRANAT II NO. 4	004	005
513340	3171033012530003	R.LUKMAN	KUNINGAN	30|12|1953	S	L	JLN HOWITZER RAYA	010	001
513341	3171036212570007	AMINI	JAKARTA	22|12|1957	S	P	JLN HOWITZER RAYA	010	001
513342	3171032606791002	HASBI SARJONO	PALEMBANG	26|06|1979	S	L	SUMUR BATU	011	001
513343	3171035001920001	SITI MAISAROH	JAKARTA	10|01|1992	S	P	SUMUR BATU	011	001
513344	3171030301011001	GALIH ADI KUNCORO	GROBOGAN	03|01|2001	B	L	JL.SUMUR BATU	011	001
513345	3325055507000008	NUR LILIS BAROKAH	PEKALONGAN	15|07|2000	B	P	SUMUR BATU RAYA	011	001
513346	3201131204000007	MUHAMAD SYEMMIL AL AZZAM	JAKARTA	12|04|2000	B	L	SUMUR BATU	011	001
513347	3201136511010005	ANNISA JIHAN FAIRUS.R	JAKARTA	25|11|2001	B	P	SUMUR BATU	011	001
513348	3171032410010004	FADHIILAH DAFFA CAHYA C	JAKARTA	24|10|2001	B	L	JL EMAS/42 SUMUR BATU	011	001
513349	3171036410010003	NABIILAH DEFFI CAHYA B	JAKARTA	24|10|2001	B	P	JL EMAS/42 SUMUR	011	001

							BATU		
513350	3171030112000003	MUHAMMAD RIZKY RACHMAN	JAKARTA	01|12|2000	B	L	JLN SUMUR BATU	011	001
513351	3171031104010007	DZIKRULLAH NUR AMIN	JAKARTA	11|04|2001	B	L	SUMUR BATU	011	001
513352	3171035112010006	DELLA FTRIA YAHYA	JAKARTA	11|12|2001	B	P	JL. SUMUR BATU	011	001
513353	3171035808000004	FELICIA JONATHAN	JAKARTA	18|08|2000	B	P	JLN HOWITZER LANJUTAN NO 26	011	001
513354	3171036810000004	VIDIA WULANDARI	JAKARTA	28|10|2000	B	P	JL. SUMUR BATU	010	001
513355	3171034112490002	NURMANI	JAKARTA	01|12|1949	P	P	JL. SUMUR BATU	011	001
513359	3171034407600007	NURLIS	SUNGAI PAKU	04|07|1960	S	P	JL. NILAM VI	013	002
513360	3310120809890001	DWI ACHMAD SOLIHIN	KLATEN	08|09|1989	S	L	JL. HOWITZER RAYA NO 13	013	002
513361	3171033101870007	RAHMAT HERMAWAN	PALEMBANG	31|01|1987	S	L	JL. SUMUR BATU A.1	013	001
513362	3171103431088002	NUR SOLEHA	JAKARTA	03|10|1988	S	P	JL. SUMUR BATU A.2	013	001
513363	3522229005860003	SUKARNO	BOJONEGORO	09|05|1986	S	L	JL. SUMUR BATU	013	001
513364	3312204608930001	AYOK ANGGRISTINA	WONOGIRI	06|08|1993	S	P	JL. SUMUR BATU A.1	013	001
513365	3206241706880008	HENDRA	TASIKMALAYA	14|08|1988	S	L	JL. SUMUR	013	001

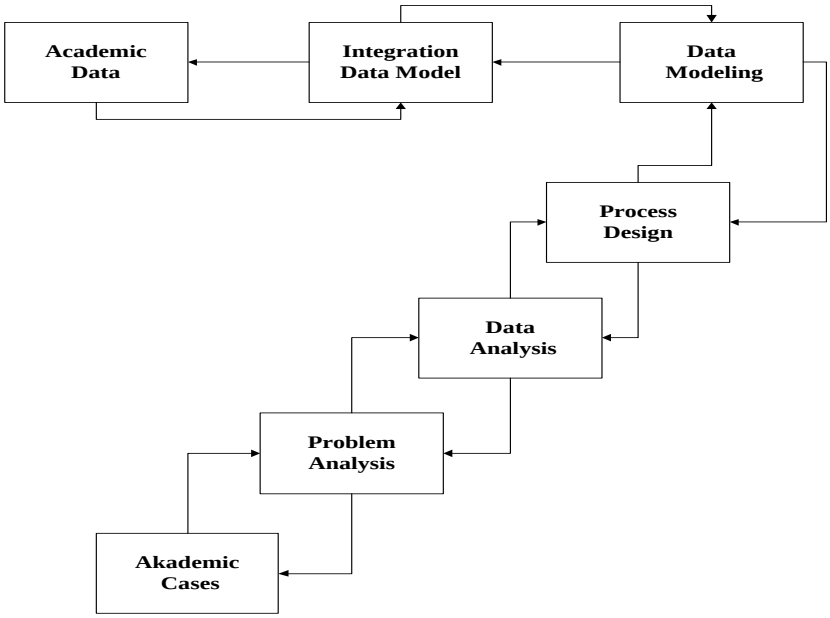
		GUNAWAN					BATU A.2		
513366	3172051510500002	SUBAGDJA	CIREBON	10|10|1990	S	L	JL. SUMUR BATU A.1	013	001
513367	3172057004760001	NURLAELY FATIMAH	JAKARTA	30|04|1976	B	P	JL. SUMUR BATU A.1	013	001
513368	3172051901810002	SOLEH NURZAMAN	JAKARTA	19|01|1981	B	L	JL. SUMUR BATU A.1	013	001
513369	3172052610850002	RAHMAT HIDAYAT	JAKARTA	26|10|1985	B	L	JL. SUMUR BATU A.1	013	001
513370	3172055803920001	YUSRINA RACHMAWATY	JAKARTA	18|03|1992	B	P	JL. SUMUR BATU A.1	013	001
513371	1304032501900005	KHADIR	SIMPANG GOBAH KP BENDANG	25|01|1990	S	L	JL. SUMUR BATU	013	001
513372	3171036808750006	JULIANA RIYANA EKAWATI	JAKARTA	28|08|1975	S	P	JL. NILAM VI NO 29	014	002
513373	3171036411610002	HARIYATIN	MALANG	24|11|1961	S	P	JL. NILAM VI NO 6B	014	002
513374	3171032209650004	SYAHRI	REMBANG	22|09|1965	S	L	JL. SUMBER PELITA GG II	005	001
513376	3171036309900008	ULIN NIMAH	REMBANG	23|09|1990	B	P	JL. SUMBER PELITA GG II	005	001
513377	3171031209940002	LUTHFIL HAKIM	REMBANG	12|09|1992	B	L	JL. SUMBER PELITA GG II	005	001
513378	3203236606960006	SITI SUNARIAH	CIANJUR	26|06|1996	S	P	JL. ROKET	003	005

							NO 11		
513379	3329136606910001	NURHAYATI	BREBES	26|06|1992	S	P	JL. MORTIR 2 NO 1	004	005
513380	3175031809820003	ROMMY SOFIAN	JAKARTA	18|09|1982	S	L	JL. MORTIR NO 1	002	005
513381	3171034103940002	SITI HASANAH	JAKARTA	01|03|1994	S	P	JL. GRANAT NO 7	002	005
513382	3275025211900011	NOVITA AJENG. P	JAKARTA	12|11|1990	S	P	JL. STEN NO 4	002	005
513383	3171036502480001	E. WARGIANTI	YOGYAKARTA	25|02|1948	S	P	JL. SUMBER PELITA	001	001
513384	3171032912810005	ANTON PUJA AKBAR	JAKARTA	29|12|1981	B	L	JL. SUMBER PELITA	001	001
513385	3171032403840003	AJIE KHAUKA PUJANTARA	JAKARTA	24|03|1984	B	L	JL. SUMBER PELITA	001	001
513386	3214046802740001	IPAH	PURWAKARTA	28|02|1974	S	P	JL. SUMBER PELITA	001	001
513387	3172025201890012	RIRIT EKA PAJARWATI	BOGOR	12|01|1989	S	P	JL. SUMBER PELITA	001	001
513388	1271044606690009	SARIANI	PADANG	06|06|1969	S	P	JL. SUMUR BATU	001	001
513389	3171035608910004	ARUMIATUN	JAKARTA	16|08|1991	S	P	JL. SUMBER PELITA	002	001
513390	3603315008860001	EVANIA	BATURAJA	10|08|1986	S	P	JL. SUMUR	002	001

							BATU		
513391	3171032812000001	JESON SALOM PAKPAHAN	JAKARTA	28|12|2000	B	L	JL.INTAN VI NO.224	006	002
513392	3171037003010009	PUTRI DWI LESTARI	JAKARTA	30|03|2001	B	P	JLN. INTAN V NO. 191	006	002
513393	3171031807000007	MUHAMMAD ALDI	JAKARTA.	18|07|2000	B	L	JLN. INTAN V	006	002
513394	3171032010000007	AKMAL FAUZAN SURANTA	JAKARTA	20|10|2000	B	L	JLN INTAN VI NO 223	006	002
513395	3301061803930003	INDRA SAPUTRA	CILACAP	18|03|1993	B	L	JL. SUMUR BATU	001	002
513396	3471909410730001	RUDY LASUT	MANADO	04|10|1973	S	L	JL. KALIMAYA 1	001	002
513397	3201042605860011	ZAENAL ABIDIN	SUKABUMI	26|05|1986	S	L	JL. KALIMAYA 1	001	002
513398	3171031808310004	ABDUL RAJAB	SOLO	18|04|1981	B	L	JL. SUMUR BATU	001	002
513399	3171031505980008	TRIYONO	JAKARTA	15|05|1998	B	L	JL. SUMUR BATU	001	002
513400	3171036809941002	FITRIA SARI	CILACAP	28|09|1994	B	P	JL. SUMUR BATU	001	002
513401	1471076205780001	THERESIANA	SEMARANG	22|05|1978	S	P	JL. KALIMAYA 1	001	002
513402	3201045012890004	MAEMUNAH	BOGOR	10|12|1989	S	P	JL. KALIMAYA 1	001	002
513403	3172027010880001	SITI ROPIKOH	JAKARTA	30|10|1988	S	P	JL.	002	002

							KALIMAYA 1		
513404	3171035509850001	ASRIDA AINI ASRIL	MATARAM	15|09|1985	B	P	JLN LANCAR II NO.14	008	007
513405	3301105205720002	TURAHMI	KEBUMEN	12|05|1972	S	P	JL. SUMUR BATU	001	007
513406	3171031404981001	ANGGA ARIN RUSWANTI	WONOGIRI	14|04|1998	B	L	JL. SUMUR BATU	001	007
513407	3175011010700016	KIKI	JAKARTA	10|10|1970	B	L	JL. SUMUR BATU	001	007
513408	3171055579300041	YENIFARIDA	JAKARTA	15|07|1993	S	P	JL. SUMUR BATU	001	007
Total Data Measured amounted to 5185									

The research method to be used in the measurement of academic data is to use the cycle model of integration of academic data of students, with the hierarchy of grid partition (HGP) method as follows:


Academic Data
Integration Data Model
Data Modeling
Process Design
Data Analysis
Problem Analysis
Akademic Cases

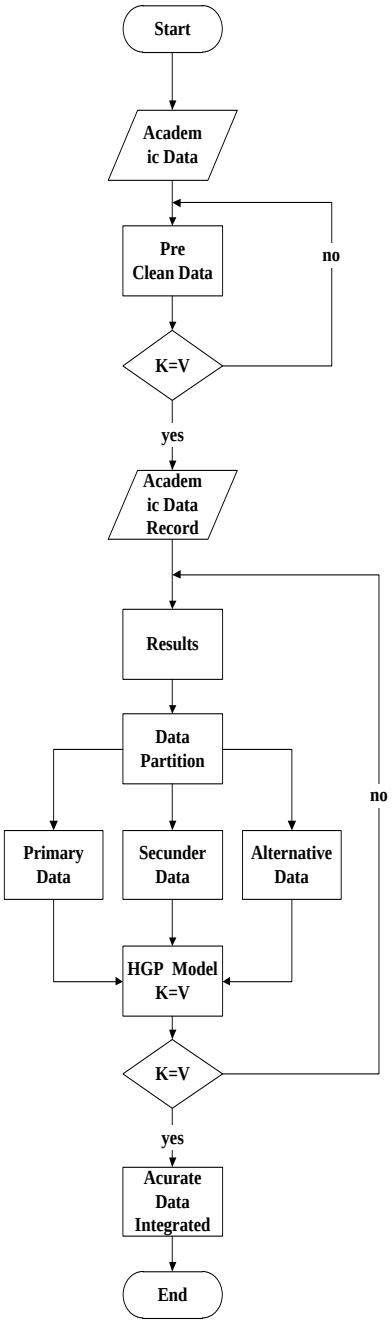


Figure 1. Research Methods – Hierarchy of Partition Grids (HGP)

The model developed in this research text, is a model of measuring the integration of student databases in optimizing academic data, by making a model determining the similarity of student data attributes in the student academic data integration system can reduce the creation of Fake Diploma Data by individuals who have certain interests / needs.

RESULTS AND DISCUSSIONS

The model developed by researchers in this Manuscript, is a form of external proof of the dissertation research results. This academic data measurement model, researchers use the Hierarchy of Partition Grid (HGP) model.

HGP Formulation Model:

K = V

K = (Vu), (Vs), (Va)

Information:

K = Size of Attribute Similarity

V = Similarity attribute variable

Vu = Main object attribute variable

Vs = Secondary object attribute variables

Va = Alternative object attribute variable

The following is the concept of the similarity attribute of student academic data, and is divided into partition variable attributes as follows:

Table 2. Data Partition

Partition 1: Main similarity attribute variable (Vu):	**Partition 2: Secondary similarity attribute variables (Vs):**	**Partition 3: Alternative similarity attribute variables (Va):**
Student ID	Certificate serial	Religion
Number	number	Dara Group
ID number	Study program	Fingerprint
Student name	College	Face
Date of birth		Home address
Gender		Kelurahan
Place of birth		sub-district

Biological mother's name		districts Province

Figure 2.Flow Data integration system process (HGP)

Moedel Process Measurement of integrated academic data, can be seen in the use case process diagram (UML) image below:

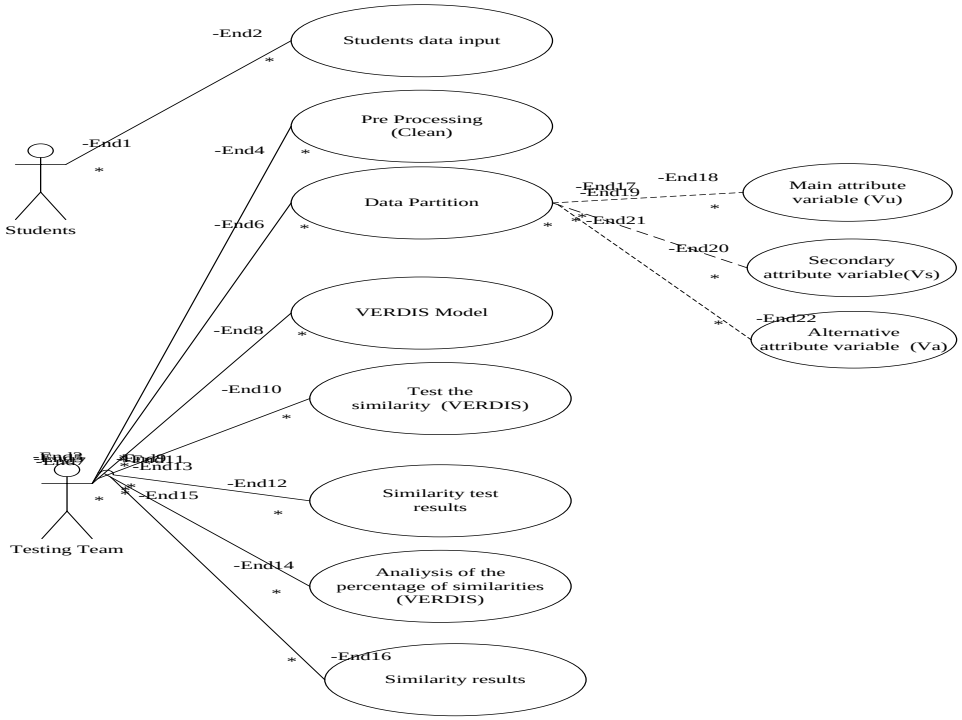


Figure 3. Use case diagram testing academic data

While the relational database model of the student academic data testing system, as follows:

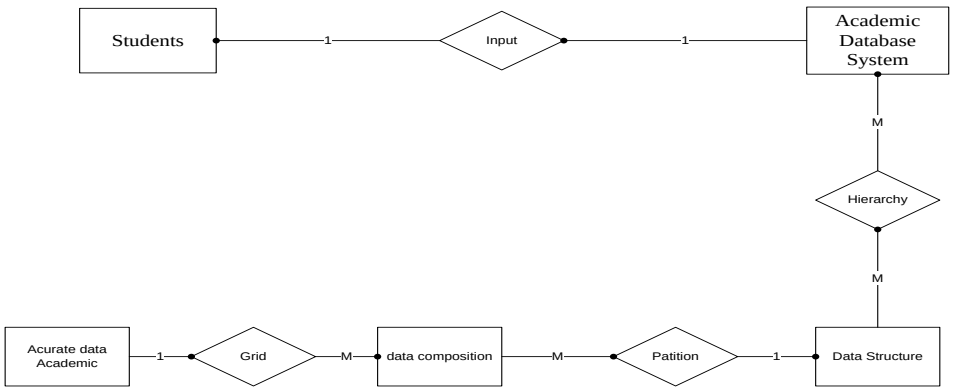


Figure 4. Entity relationship diagram testing students data

Data Integration Formulation Model in the database, HGP Model for the primary similarity attribute variable and the secondary similarity attribute variable, as follows:

K = V

K = (Vu), (Vs) ,(Va)

K = (Student Identification Number,Resident Identification Number, Student Name, Date of Birth, Gender,Place of Birth,Birth Mother's Name), (Academic Serial Number,Study Program,Higher Education),(Religion,Blood Type,Face, Fingerprint,Residential Address,Village,District,District,Province)

1. Test resu lt

In this study, researchers analyzed thousands of data to be carried out the process of data integration testing, as proof of the data testing. The following is a display of testing data in a database and displaying the results of the test.

The following figure is the layout of the data tested in the database, Indonesian instructions and the following figure is a form of data similarity testing in a SQL database

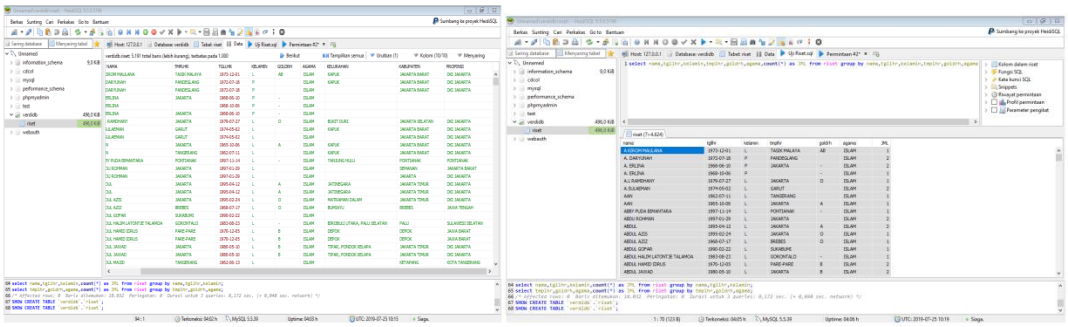


Figure 5. Data in SQL database.Figure 6. Model Testing Data similarity in SQL databases.

From the process of measuring the accuracy of the similarity of population data to be adjusted to academic data, using the Hierarchy of Grid Partition (HGP) model, then the results can be seen in the table of similarity measurement results below:

Table 3.Table Of Results Measurement of similarity of population and academic data

Similarity index category (Ki)	Total attributes similar	Percentage Similarity analysis	Similarity analysis results
Index Category = 4	4	0,08%	Fraud or Similar
Index Category = 3	138	2,66%	Similar (Invalid)
Index Category = 2	932	17,97%	Similar (Invalid)
Index Category = 1	4111	79,29%	Single Indentity
Total Data	5185	100,00%	

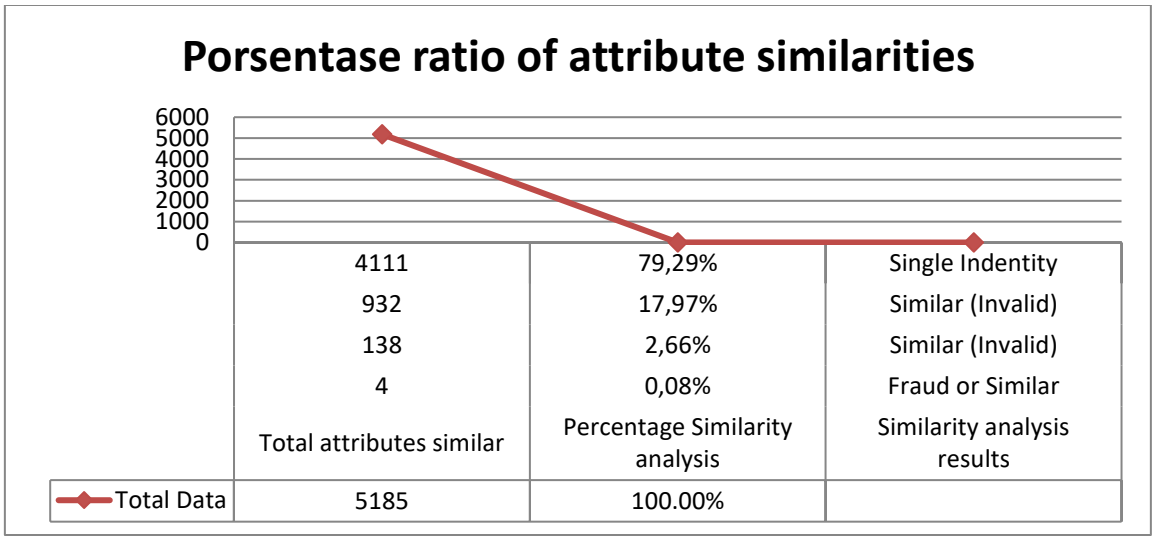


Figure 7.Measurement of the similarity of population data attributes for student academic data using SQL databases.

CONCLUSION

Measuring the accuracy of student academic data with a hierarchical model of partition grid (HGP) integrated database system, certainly has an impact on effectiveness in the process of finding academic data of students and alumni, so that the data found is accurate student data, according to the actual data and concept of this model, reducing the potential for anyone who makes falsification of academic data identity.In order to build a database system that can maintain student integrity. Furthermore, this paper will be developed in the following research process, both by the researchers themselves and by other researchers, which is relevant to the topic of this research. This research text is not perfect,But the measurement model carried out in researchers can be done in testing / measuring data accuracy by integrating student academic data with population data, so that it is easy to access and find the truth of student data.The results of measuring the accuracy of data similarity using the Hierarchy of Grid Partition (HGP) population data for student academic data, The results of measuring the accuracy of data similarity using the Hierarchy of Grid Partition (HGP) population data for student academic data, from a total of 5185 data, it can be seen that the measurement results that amounted to 4111 normal data (single ID), amounted to 1070 invalid double data (similarities), and amounted to 4 valid double data (similarities), so it can be concluded that there are 4 data that have a valid level of accuracy of similarity.

ACKNOWLEDGMENT

The data presented in this research text, is part of the research data dissertation of the Doctoral Program in Computer Science, which discusses the measurement of data equality, based on data sets of student population, measured for academic data. This research was accompanied by Professor Dr. Muhammad Zarlis, Prof. Dr. OpimSalimSitompul, PoltakSihombing, PhD.

REFERENCE

[1] ArshiGouhar, 2017, Database Management System, "International Journal of Engineering Science and Computing", May 2017, Volume 7, Issue 5 (2017), Page11766-11768, http://ijesc.org/

[2] RadoslavaKraleva, VelinKralev, Nin! Sinyagina, 2018, Design and Analysis of a Relational Database forBehavioral Experiments Data Processing,"International Journal of Online Engineering (iJOE) · February 2018 ",Vol. 14, No. 2, 2018, Page 117-132, https://doi.org/10.3991/ijoe.v14i02.7988URL:http://www.i-joe.org

[3] AzharSusanto, Meiryani, 2019, Database Management System, "INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH ", VOLUME 8, ISSUE 06, JUNE 2019, ISSN 2277-8616, Page 309-312, IJSTR©2019 www.ijstr.org

[4] Kristi L. Berg,Tom Seymour,RichaGoel, 2013,History Of Databases, "International Journal of Management & Information Systems (IJMIS) · December 2012Volume 17, Number 1 (2013), Page 29-36, © 2013 The Clute Institutehttp://www.cluteinstitute.com/

[5] AbdulrahmanHamedAlmutairi&AbdulrahmanHelalAlruwaili, Security in Database Systems, "Global Journal of Computer Science and Technology Network, Web & Security", Online ISSN: 0975-4172 & Print ISSN: 0975-4350,Volume 12 Issue 17 Version 1.0 Year 2012, Page 8-14.

[6] Wang, N., Chen, X., Song, G., Parsaei, H. (2015). An experiment scheduler and federated authentication solution for remote laboratory access. International Journal of Online Engineering, 11(3): 20-26. https://doi.org/10.3991/ijoe.v11i3.4554

[7] Limpraptono, F. Y., Ratna, A. A. P., Sudibyo, H. (2013). New architecture of remote laboratories multiuser based on embedded web server. International Journal of Online Engineering, 9(6): 4-11. https://doi.org/10.3991/ijoe.v9i6.2886

[8] Wagner, J., Rasin, A., Glavic, B., Heart, K., Furst, J., Bressan, L., Grier, J. (2017). Carving database storage to detect and trace security breaches. Digital Investigation, 22: 127-136. https://doi.org/10.1016/j.diin.2017.06.006

[9] Kim, J., Park, A., Lee, S. (2016). Recovery method of deleted records and tables from ESE database. Digital Investigation, 18: 118-124. https://doi.org/10.1016/j.diin.2016.04.003

[5] Wagner, J., Rasin, A., Grier, J. (2016). Database image content explorer: Carving data that does not officially exist. Digital Investigation, 18: 97-107. https://doi.org/10.1016/j.diin.2016.04.015

[10] Tosun, U. (2014). Distributed Database Design: A Case Study.Procedia Computer Science, 37: 447-450. https://doi.org/10.1016/j.procs.2014.08.067

[11] Currim, S., Ram, S., Durcikova, A., Currim, F. (2014). Using a knowledge learning framework to predict errors in database design. Information Systems, 40: 11-31. https://doi.org/10.1016/j.is.2013.08.001

[12] Nicolaos, P., Katerina, T. (2015). Simple-talking database development: Let the end-user design a relational schema by using simple words. Computers in Human Behavior, 48: 273-289. https://doi.org/10.1016/j.chb.2015.02.002

[13] Gottlob, G., Pichler, R., Wei, F. (2010). Tractable database design and datalog abduction through bounded treewidth. Information Systems, 35(3): 278-298. https://doi.org/10.1016/j.is.2009.09.003

[14] Osman, R., Awan, I., Woodward, M. E. (2011). QuePED: Revisiting queueing networks for the performance evaluation of database designs. Simulation Modelling Practice and Theory, 19(1): 251-270. https://doi.org/10.1016/j.simpat.2010.06.010

[15] Dimitrieski, V., Celikovic, M., Aleksic, S., Ristic, S., Alargt, A., Lukovic, I. (2015). Concepts and evaluation of the extended entity-relationship approach to database design in a multi-paradigm information system modeling tool. Computer Languages, Systems & Structures, 44(C): 299-318. https://doi.org/10.1016/j.cl.2015.08.011

[16] Halpin T, Morgan T, 2008, Information Modeling and Relational Database, 2008. ISBN 978-0-12-373568-3. (In indonesian)

[17] Fowler, Martin, 1997, Pattern Analysis, reusable object models, Addison-Wesley Longman, ISBN 0-201-89542-0.

[18] Stewart Robinson, Roger Brooks, Kathy Kotiadis, and Durk-Jouke Van Der Zee (Eds.): 2010, Conceptual Modeling for Discrete Event Simulation. ISBN 978-1-4398-1037-8

[19] David W. Embley, Bernhard Thalheim (Eds.), 2011, Conceptual Modeling Handles. ISBN 978-3-642-15864-3.

[20] Verdi Yasin, 2012, RekayasaPerangkatLunakBerorientasiObjek (Object-Oriented Software Engineering): Modeling, Architecture and Design (In Indonesian), Publisher Mitra Wana Media, Jakarta – Indonesia

[21] Abraham Silberschatz and Henry F. Korth and S. Sudarshan, 2011, Database System Concepts, sixth edition, McGraw-Hill Companies, Inc., ISBN 978-0-07-352332-3

[22] David L. Olson and Dursun Delen, 2008, Advanced Data Mining Techniques, Springer-Verlag Berlin Heidelberg, ISBN: 978-3-540-76916-3 e-ISBN: 978-3-540-76917-0

[23] David L. Olsen, 2007, Big Data, http://www.gartner.com/it-glossary/big-data (access 10 June 2018)

[24] Djoni Setiawan K., and Hendra Bunyamin, 2018, Utilization of inverted index in the process of tracing the contents of pdf document documents of students' final assignments, National Seminar on Information and Communication Technology 2018 (SENTIKA 2018) ISSN: 2089-9815, Yogyakarta, March 23 2018 page 356 -365

[25] Felipe Gómez-Cuba and Rafael Asorey-Cacheda and Francisco J. González-Castañ, 2013, Smart Grid Last-Mile Communications Model and Its Application to the Study of Leased Broadband Wired Access, IEEE Transactions on Smart Grid, Vol. 4, No. 1, March 2013

[26] Gehlot et al., 2012, International Journal of Modern Engineering Research (IJMER) www.ijmer.com Vol.2, Issue.4, July-Aug 2012 pp-1923-1928 ISSN: 2249-6645.

[27] González-Castaño. IEEE Transactions on smart grid, vol. 4, no. 1, March 2013

[28] Humasak Simanjuntak, Rosni Lumbantoruan, Wiwin Banjarnahor, Erisha Sitorus, Magdalena Panjaitan, Sintong Panjaitan, 2017, Similarity Assessment Entity Relationship Diagram with Tree Algorithm Edit Distance, JNTETI, Vol. 6, No. February 1, 2017

[29] H. Edelsbrunner, Z. Gu, 2008, Design and Analysis of Algorithms, Durham: Duke University

[30] Verdi Yasin, Muhammad Zarlis, OpimSalimSitompul, PoltakSihombing. 2020, Hierarchy of Grid Partition (HGP) Integrating data inSoftware Engineering and databases, “International Conference on Applied Sciences, Information and Technology 2019”, IOP Conf. Series: Materials Science and Engineering **846** (2020) 012024 , IOP Publishing, doi:10.1088/1757-899X/846/1/012024 .

[31] Verdi Yasin, Muhammad Zarlis, OpimSalimSitompul, PoltakSihombing, 2019, Big data measurement model in achievingmaximum accuracy using the model Hierarchy ofGrid Partition (HGP) method, “2019 The 3rd International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM)”, 978-1-7281-2475-9/19/$31.00 ©2019 IEEE, Page 107-110.

[32] Rumbaugh, James; Jacobson, Ivar; Booch, Grady. (1999) " The Unified Modeling Language Reference Manual (UML)" ISBN: 0-201-30998-X. Publisher: Addision Wesley Longman, Inc.First Printing, December 1998.

[33] Booch, Grady. (1994), "Object-oriented Analysis and Design with Application, 2nd edition. ISBN: 0-8053-5340-2. Publishers: Addision Wesley Longman, Inc. First Printing, December 1998.

[34] Sommerville, Ian. (2003), "Software Engineering",ISBN: 979-688-946-3 Volume 1, 6th edition (in Indonesian) Jakarta-Indonesia, Publishers: Erlangga.