# Business Intelligence for Unemployment Rate Management System

by Fathoni Fathoni

Submission date: 09-May-2023 01:25PM (UTC+0700)

**Submission ID:** 2088348073

**File name:** 3128-7830-1-PB.pdf (767.5K)

Word count: 3951

Character count: 21558

### The IJICS (International Journal of Informatics and Computer Science)

Vol 5 No 2, July 2021, Page 111-117

ISSN 2548-8384 (online), ISSN 2548-8449 (print)

Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/ijics/index DOI 10.30865/ijics.v5i2.3128



#### **Business Intelligence for Unemployment Rate Management System**

Tika Octri Dieni, Ken Ditha Tania\*, Fathoni, M Ihsan Jambak, Pacu Putra

Departement of Information System, University of Sriwijaya, Palembang, Indonesia Email: 1tikaoctridieni@gmail.com, 2.\*kenya.tania@gmail.com, 3Fathoni@Unsri.ac.id, 4Jambak@unsri.ac.id, 5Pacuputra@ilkom.unsri.ac.id Coressponding author: kenya.tania@gmail.com

Submitted: 07/07/2021; Accepted: 14/07/2021; Published: 31/07/2021

Abstract-Unemployment rate is one of many problems which being faced by the government for each country, especially in Indonesia. Based on statistics data, Central Bureau of Statistics (Indonesia), unemployment rate in Indonesia is quite high which scattered throughout the province. Those data was obtained with a long process, by the time, human resources and cost which are not small. Thus, this paper proposes business intelligence for unemployment rate management system to predict from its several causes which are potentially on increasing unemployment rate, using Business Intelligence Roadmap methods, it used because its adaptive and detail which consists of 18 stages from 6 phases. There is one of the 18 stages, namely data mining, for this data mining using the KNN algorithm. Business intelligence can processing data into useful information or knowledge. Generally, business intelligence has main process which are data collection, then this data would be processed by ETL (Extract, Transform, Load) before get into data warehouse as a place data storage, so that those data could functionally used for analysis process with OLAP and data mining to classify the result of unemployment rate prediction from its several potentially causes. This paper would possibly to find patterns in the unemployment rate and its causes then the result of the pattern will visualizations on web application with business intelligence based that would be developed that easy to used / user friendly and attractive user interface.

Keywords: Unemployment Rate; Business Intelligence; Business Intelligence Roadmap; KNN Algorithm; ETL; OLAP

#### 1. INTRODUCTION

Generally Business Intelligence (BI) used for those who have business or for company as a tool for processing transactional data which they have to change it to be an usefull information and not only fixated on the numbers so as to be able to provide a strategy in producing a decision. Data is the valuable asset for many company also agency. In fact, The flow of information is an important factor in achieving the smooth running of a company. Now many companies collect and store large amounts of information that is processed from various data from year to year, which is the basis of Business Intelligence [1]. Business Intelligence is often applied to the private sector with the aim of making a profit, it is also applied to the public sector or government agencies. As stated by [2] "BI has been widely applied to the private or private sector which in the end yielded good results, therefore BI has good potential in other fields such as public services and non-profit organizations driven by the government".

The main of Business intelligence process are data collection, data storage, and data use [3]. At first is data collection, then this data would be processed by ETL (Extract, Transform, Load) before get into data warehouse as a place data storage, so that those data could functionally used for analysis process with OLAP and data mining to classify the result of unemployment rate prediction from its several potentially causes. This paper would possibly to find patterns in the unemployment rate and its causes then the result of the pattern will visualizations on web application with business intelligence based that would be developed that easy to used / user friendly and attractive user interface.

In Business Intelligence there is a process called ETL (Extract, Transform, and Load) that can extract data from various sources, eliminating noise data and changing the data format suitable with the design that has been made, then it is entered into the data warehouse (DW) [4]. On the process of making data warehouse (DW) to acommodate data so that it can be analyzed then displayed and produce useful information that sourced from the data [3]. The data contained in the data warehouse will be analyzed and visualized [5] also processed by data mining using the classification method using K-Nearest Neighbor. This approach is an algorithm used to create new knowledge in the form of predictions and must initialize the value of K.

The Central Bureau of Statistics or BPS can be abbreviated as the role of providing data facilities to those in need, such as the government and the community. The data obtained are obtained through censuses and surveys of the population as well as access through related departments and institutions as secondary data. Based on data published by BPS entitled Population Projections by Province, 2010-2035, it is stated that Indonesia has a population of 269,603,400 people in 2020 spread throughout the Indonesian Province (bps.go.id). This certainly affects all aspects of life such as the aspect of employment.

Based on data from the Central Statistics Agency in 2019, Indonesia has a workforce consisting of 129,366,192 people who work and 6,816,840 unemployed, which means it is still quite high. The data at BPS is obtained through a survey or population census, which means that this method is considered ineffective and inefficient. For this reason, it is necessary to have a technique in computer science, namely the application of Business Intelligence that can process data into useful new 4 formation or knowledge. This can be a solution to help analyze of unemployment rate from various causes through the design and development of web applications with Business Intelligence based. This platform will analyze several potential causes of increasing unemployment rates and also make visualizations that we can see by graphic or diagram on web.



Ther 4 are several research studies that related to business intelligence, such as in research [3] at health institutions, developing a pervasive web application based on business intelligence (BI) clinical indicators in an attempt to reduce the number of appointments, surgeries, and medical examinations that were not carried out in a Portuguese health institution most likely due to forgetfulness. Implementation of BI in the health sector [6][7][8], because it can analyze and provide data on the health care system: high quality, responsive, affordable and efficient and BI indeed proper way to transform the office shelves to well-automated system.

BI currently applied in small 41 medium-size companies to helping in managing amount data. Then regarding absenteeism rates on company [9], it consisted in designing and developing a web applicati 4 based on Business Intelligence (BI) Indicators to identified the causes for an increasing absenteism rate and analyzed of the relation between the great number of absences and the health also lifestyle of collabilitators. In research [10][11] within the bank, the existence of BI can help provide decisions in the form of policies to customers to continue to improve transactions and bank management reporting facilities. In the public sector [2][12][13], BI helps the management and analysis of evergrowing data into useful information 3 ith a more effective process and productive environment. In education (higher education) [14][15][16], BI is needed to improve the quality of real-time decision making.

Base 4 on [9], Business intelligence concept refers to collecting, transforming, organizing, analysing, and distributing data from multiple internal and external sources of information for a faster and more effective decision-making process. BI can transforms a large amount of data into valuable information/knowledge, which can provides the possibility of taking and making fact-based decisions which are more strategic and effective.

#### 2. RESEARCH METHODOLOGY

#### 21 Method of Collecting Data

There are several methods that used on collecting process of data and information on this studi, as follows.

- a. Observation
- b. Literature review
- c. Ingrview

In conducting this research,the sources of data obtained are from the leader, employees, and the official website of the Central Bureau 1 Statistics.

There are two types of data which used in this study, primary data and secondary data. Data primer are those obtained directly from the source or object of research. Primary data in this study were obtained through direct interviews with informans. Meanwhile, secondary data is a data that has been published or has been used by other parties. Secondary data are supporting data in this study obtained from institutions/agencies involved in this research, including the Central Bureau of Statistics.

#### 2.2 System Development Method

In this study, the method that will be used and discussed is using the Business Intelligence Roadmap approach [17] as the following picture.

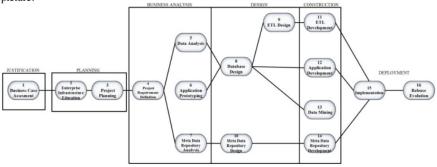


Figure 1. Business Intelligence Roadmap [17]

The explanation of the stages in the business intelligence project life cycle method is as follows.

a. Justification Stage

In this justification stage, a business case assessment will be carried out, where a business case assessment is the first step that is considered for those who will develop BI, as in this study, namely the identification of things needed in BI development at the Central Bureau of Statistics. The activities carried out at this stage are:

- 1. Evaluating data sources obtained and owned by the Central Bureau of Statistics.
- 2. Identifying problems and opportunities at the Central Bureau of Statistics.
- 3. Determine business needs at the Central Bureau of Statistics.
- 4. Determine the objectives or goals of the Business Intelligence application.

Vol 5 No 2, July 2021, Page 111-117

Tika Octri Dieni, Business Intelligence for Unemployment Rate Management System

5. Propose a Business Intelligence solution.

#### b. Planning Stage

At the planning stage there are two main activities, namely:

1. Enterprise Infrastructure Evaluation

In the enterprise 3 frastructure evaluation stage, an evaluation of the technical infrastructure will be carried out which includes hardware, software, middle-ware, database n alagement systems, operating systems, network components, meta data repositories, and others, as well as on non-technical infrastructure, which includes meta data standards, data-naming standards, enterprise logical data models, methods, instructions, testing procedures, and others.

2. Project Planning.

BI projects are dynthic so any changes that occur in the scope, staff, budget, technology, business processes can have an impact on the success of a BI project. Therefore, the project planning must be made in more detail and the latest progress must always be monitored and reported.

#### c. Business Analysis Stage

At the business analysis stage there are 4 main activities that are adjusted to the data obtained from Central Bureau of Statistics, namely:

1. Boject requirement definition

The activities carried out at this stage are defining the need for technical and non-technical infrastructure improvements, reports, data sources, reviewing the project scope, expanding the logical data model, defining the initial service level agreement, and writing application requirements documents according to the case.

2. Data analysis

The activities carried out at this stage are analyzing external data sources, redefining the logical data model, analyzing the quality of data sources, expanding the enterprise logical data model, correcting data discrepancies, and writing data-cleansing specifications.

	2008 Februari									
Pendidikan Tertinggi yang Difamatkan	Angkafan Kerja (AK)				Bukan Angkafan Kerja (BAK)				Jumlah Penduduk Usia 15	Persentase Angkatan Kerja
	Bekerja	Pengangguran	Jumlah AK	% Bekerja / AK	Sekolah	Mengurus	Lainnya	Jumlah BAK	tahun ke Atas	Terhadap Penduduk Usia Kerja (TPAK)
Tidak/belum pernah sekolah	5,670,322	79,764	5,750,086	98.61	-	2,568,957	1,599,878	4,168,835	9,918,921	57.
Tidak/belum tamat SD	12,982,233	448,431	13,430,664	96.66	95,968	4,263,968	1,429,761	5,789,697	19,220,361	69.
SD	36,963,023	2,216,748	39,179,771	94.34	3,154,334	11,592,816	2,589,403	17,336,553	56,516,324	69.
SLTP	19,396,319	2,166,619	21,562,938	89.95	7,252,367	6,693,408	1,240,744	15,186,519	36,749,457	58.
SLTA Umum/SMU	13,899,839	2,204,377	16,104,216	86.31	2,272,737	4,265,402	1,112,755	7,650,894	23,755,110	67.
SLTA Kejuruan/SMK	6,708,513	1,165,582	7,874,095	85.20	341,274	1,631,078	328,060	2,300,412	10,174,507	77
Akademi/Diploma	2,660,606	519,867	3,180,473	83.65	121,589	630,005	203,687	955,281	4,135,754	76
Universitas	3,769,002	626,202	4,395,204	85.75	42,838	477,135	180,381	700,354	5,095,558	86.
Tak Terjawab	-	-	-	-	-	-	-	-	-	
Total	102,049,857	9,427,590	111,477,447	91.54	13,281,107	32,122,769	8,684,669	54,088,545	165,565,992	67
	2009 Februari									
Pendidikan Tertinggi yang Ditamatkan	Angkatan Kerja (AK)				Bukan Angkatan Kerja (BAK)				Jumlah Penduduk Usia 15	Persentase Angkatan Kerja
	Bekerja	Pengangguran	Jumlah AK	% Bekerja / AK	Sekolah	Mengurus	Lainnya	Jumlah BAK	tahun ke Atas	Terhadap Penduduk Usia Kerja (TPAK)
Tidak/belum pernah sekolah	5,698,092	60,347	5,758,439	98.95		2,726,990	1,848,511	4,575,501	10,333,940	56
Tidak/belum tamat SD	13,487,937	415,955	13,903,892	97.01	160,685	6,463,881	2,250,469	8,875,035	22,778,927	61
SD	36,243,801	2,143,747	38,387,548	94.42	3,284,146	9,468,665	1,674,068	14,426,879	52,814,427	72
SLTP	19,846,539	2,054,682	21,901,221	90.62	7,623,326	6,226,496	1,117,069	14,966,891	36,868,112	59
SLTA Umum/SMU	15.131.145	2 133 627	17.264.772	87.64	2 227 954	4.188.247	733.614	7.149.815	24.414.587	70

Figure 2. Example of Data on Microsoft Excel

#### 3. Spplication prototyping

The activities carried out at this stage are analyzing access needs, analyzing the scope of the prototype charter, designing reports and queries, building prototypes, and demonstrating prototypes.

#### 43 Meta data respiratory analysis

The activities carried out at this stage are analyzing the meta data repository requirements, analyzing the interface requirements for the meta data repository, analyzing the meta data repository access and reporting requirements, creating a logical meta model, and creating meta-data.

#### d. Design Stage

At the design stage there are 3 main activities that are adjusted to the data obtained from Central Bureau of Statistics, namely:

#### 1. Database design

Database design is a medium for storing large amounts of data on a regular basis that will be used in the development of Business Intelligence systems. The database design process on the BI system is carried out in accordance with the needs analysis that has been made previously. BI database design which will be designed in a multidimensional database.

Tika Octri Dieni, Business Intelligence for Unemployment Rate Management System

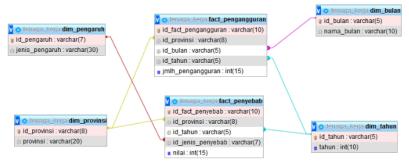


Figure 3. Star Schema

#### 2. Extract, Transform, Load (ETL) Design

The quality and format of the data source is not good, such as noise and null values, it will take a lot of time for transformation and cleaning, so it requires an ETL design so that the data used is clean and also to get the same data and format.

#### 3. Metadata Repository Design

The design phase of the metadata repository functions to design a logical data model documentation in the data warehouse in detail about the data in the data warehouse which includes field names, field types, and field sizes as well as the processes involved.

#### e. Construction Stage

At the design stage there are 3 main activities that are adjusted to the data obtained from Central Bureau of Statistics, namely:

- Extract, Transform, Load (ETL) Development
   Its activities include creating and processing ETL, integrating the ETL process, ETL process performance, ETL quality assurance process, and ETL acceptance process. In this process using pentaho spoon tools.
- 2. Application Development

BI system application development for cube creation aims to display data analysis that processed using pentaho noon tools through the website platform.

3. Data Mining

Mining or discovering of new information by looking for certain patterns or rules from a very large amount of ta. In this study, the data mining process uses the KNN algorithm.

Meta data repository development

The results of the design for the meta data repository carried out in the previous stage are then made.

#### f. Deployment Stage

After passing the development stage, the next step is to carry out the deployment process. At this stage there are 2 main activities that are adjusted to the data obtained from Central Bureau of Statistics, namely:

1. Implementation

The process of implementing Business Intelligence applications which includes adjustments to application users from the tools used and the users who will use them.

2. Release Evaluation

All previously released processes, techniques, instructions, etc. are studied and evaluated to produce better and more efficient applications.

#### 2.3 KNN Method

K-Nearest Neighbour (KNN) is an supervised learning, which means classification method, that stores all available cases and classifies new cases based on similirity measure [18]. KNN algorithm has an advantage, technically the operation of KNN is simple and easy therefore the content and process is easy to read and understand [19]. We need to define the k value, so it will classified based on the nearest k samples in the KNN algorithm [19]. Below is the formula of KNN:

$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^{n} (x_{1i} - x_{2i})^2}$$
 (1)

Where:

 $dis(x_1, x_2) = distance$  between two vector  $x_1$  and  $x_2$ 

 $x_{1i}$  = testing data

 $x_{2i}$  = sample data

Tika Octri Dieni, Business Intelligence for Unemployment Rate Management System

#### 3. RESULT AND DISCUSSION

By applying Business Intelligence to the Central Bureau of Statistics, it can be concluded that the BI system can help employees, especially those who hold policies related to the unemployment rate in every province of Indonesia. In terms of knowing the unemployment rate, it can be made easier by using the BI system, which is sufficient by collecting previous data which is then processed through the ETL (Extract, Transform, Load) process using Pentaho data integration tools with the aim of eliminating bad data such as still valuable data. null, this noise is adjusted in the desired data format so that the data is ready to proceed to the data analysis process.



Figure 4. ETL Process

If you have passed all the stages in ETL development, the next stage is making cube on multidimensional data for OLAP with the help of schema workbench tools. This tool can be published and saved with Pentaho dashboard analysis.

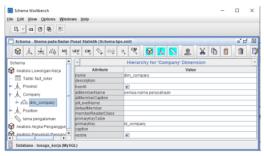


Figure 5. Result of Schema Workbench Process

Then, after the ETL and OLAP stages are carried out, the next stage is well site development. The development of a Business Intelligence tools to monitoring the unemployment rate in Indonesia is livided into two main pages based on the role of users who interact with the system, namely company leaders and admins, where each user has a menu according to their respective responsibilities.



Figure 6. Interface of Web Application Program with BI Development Concept

The following shows the results of the OLAP analysis. The information possessed regarding the unemployment rate, this analysis involves the dimensions of the province, month, year, and the number of unemployment figures which can be seen in slices and dices.



Figure 7. Analysis of Unemployment Rate

Vol 5 No 2, July 2021, Page 111-117

Tika Octri Dieni, Business Intelligence for Unemployment Rate Management System

In this data mining process using KNN algorithm with K initiated by 3 and the process flow can be seen in the image below.

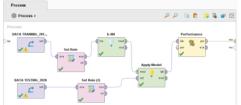


Figure 8. Data Mining Process Flow on RapidMiner

The following are the results of the data mining process obtained using the classification method with the K-Nearest Neighbor (KNN) Algorithm. So that the results in the form of new knowledge include information on predictions or forecasters, unemployment data can be said to be "reasonable" or "unreasonable".

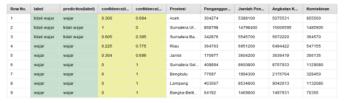


Figure 9. Classification Result Data Using KNN

From the following figure, it can be seen the performance generated using the KNN classification. Based on the processed data, an accuracy of 88.24% is obtained, which is obtained from precision and recall.



Figure 10. Result of Performance Measures of KNN Classification

#### 4. CONCLUSION

Based on the analysis which carried out and the results of the discussions that have been described in previous chapters regarding Business Intelligence system are as follows. In this study, the data used are data sourced directly from the field related to the unemployment rate. The data consists of the unemployment rate, population, poverty rate, and labor force figures, each of which is obtained from the Central Bureau of Statistics. This study helps to analyze the unemployment rate from several aspects that cause this which can be seen through a system developed with website-based Business Intelligence so as to obtain information that is more effective than using the previous tools (Ms. Word). The steps carried out in this research use the Business Intelligence Roadmap. In the construction stage process, the ETL development section is assisted by using Pentaho Data Integration tools and Schema Workbench as well as the BI server which in the end gets the information results which will be visualized through the website as well as the data mining process section with the help of RapidMiner tools. The application of data mining using the classification method with KNN with K initiated by 3 can be said to be good because it produces a prediction with an accuracy of 88.24%. The series of processes carried out in this study produce new knowledge that can assist policy holders in terms of unemployment rates in the decision-making process. The resulting knowledge is displayed in the form of a graph that is more attractive and comfortable to see. In this study, the prediction of the unemployment rate is whether it will be said to be reasonable or not.

#### REFERENCES

- [1] M. O. Nketia, "THE INTERNATIONAL JOURNAL OF BUSINESS & MANAGEMENT Business Intelligence (BI); A System to Optimize Business Performance," vol. 4, no. 9, pp. 218–224, 2016.
- [2] I. Wowczko, "Business Intelligence in Government Driven Environment," Int. J. Infonomics, vol. 9, no. 1, pp. 1107–1111, 2016.
- [3] M. Esteves, A. Abelha, and J. Machado, "The development of a pervasive Web application to alert patients based on business

- intelligence clinical indicators: a case study in a health institution," Wirel. Networks, vol. 6, 2019.
- [4] Adnan, A. A. Ilham, and S. Usman, "Performance analysis of extract, transform, load (ETL) in apache Hadoop atop NAS storage using ISCSI," Proc. 2017 4th Int. Conf. Comput. Appl. Inf. Process. Technol. CAIPT 2017, vol. 2018-Janua, pp. 1–5, 2018.
- [5] A. A. Yulianto and Y. Kasahara, "Implementation of Business Intelligence with Improved Data-Driven Decision-Making Approach," Proc. - 2018 7th Int. Congr. Adv. Appl. Informatics, IIAI-AAI 2018, pp. 966–967, 2018.
- [6] R. Gaardboe, T. Nyvang, and N. Sandalgaard, "Business Intelligence Success applied to Healthcare Information Systems," Procedia Comput. Sci., vol. 121, pp. 483–490, 2017.
- [7] S. S. Ally and N. Khan, "Data warehouse and BI to catalize information use in health sector for decision making: A case study," Proc. - 2016 Int. Conf. Comput. Sci. Comput. Intell. CSCI 2016, pp. 92–97, 2017.
- [8] L. Gastaldi et al., "Measuring the maturity of business intelligence in healthcare: Supporting the development of a roadmap toward precision medicine within ISMETT hospital," Technol. Forecast. Soc. Change, vol. 128, no. November 2016, pp. 84– 103, 2018.
- [9] S. Oliveira, M. Esteves, R. Cernadas, A. Abelha, and J. Machado, "The development of a business intelligence web application to support the decision-making process regarding absenteeism in the workplace," Adv. Intell. Syst. Comput., vol. 1137 AISC, pp. 104–113, 2020.
- [10] D. Prayitno, "Application of Business Intelligence for Banking Performance Based on Products Analysis," Int. J. Progress. Sci. Technol., vol. 6, no. 2, pp. 554–569, 2018.
- [11] B. Tuccaroglu and M. Nat, "The readiness of banks for the application of Business Intelligence solutions," 13th HONET-ICT Int. Symp. Smart MicroGrids Sustain. Energy Sources Enabled by Photonics IoT Sensors, HONET-ICT 2016, no. 2, pp. 127–132, 2016.
- [12] Y. Jadi and L. Jie, "An implementation framework of business intelligence in e-government systems for developing countries: Case study: Morocco e-government system," Int. Conf. Inf. Soc. i-Society 2017, vol. 2018-Janua, pp. 138–142, 2018.
- [13] B. Oumkaltoum, E. I. Mohammed, E. B. M. Mahmoud, and E. B. Omar, "Business intelligence and EDA based architecture for interoperability of E-Government data services," 5th IEEE Int. Smart Cities Conf. ISC2 2019, pp. 402–407, 2019.
- [14] A. Sorour, A. S. Atkins, C. F. Stanier, and F. D. Alharbi, "The role of business intelligence and analytics in higher education quality: A proposed architecture," 2019 Int. Conf. Adv. Emerg. Comput. Technol. AECT 2019, 2020.
- [15] A. S. Girsang, D. A. Sunarna, A. Syaikhoni, and A. Ariyadi, "Business Intelligence for Education Management System," 2019 Int. Conf. Comput. Sci. Inf. Technol. ICoSNIKOM 2019, 2019.
- [16] J. Fowler, "Business intelligence at the university," Proc. 6th Annu. Conf. Comput. Sci. Comput. Intell. CSCI 2019, pp. 821–825, 2019.
- [17] S. S. Wiradarma, K. D. Tania, and D. Y. Hardiyanti, "Implementation of Business Intelligence in Product Services a Banking (Case Study: PT Bank Sumsel Babel Baturaja Branch)," JSI J. Sist. Inf., vol. 9, no. 2, pp. 1242–1247, 2017.
- [18] H. Y. Riskiawan, T. D. Puspitasari, F. I. Hasanah, N. D. Wahyono, and M. Fatoni Kurnianto, "Identifying Cocoa ripeness using K-Nearest Neighbor (KNN) Method," Proc. - 2018 Int. Conf. Appl. Sci. Technol. iCAST 2018, pp. 354–357, 2018.
- [19] S. Wang, R. Ma, Y. Li, and Q. Wang, "A bluetooth location method based on kNN algorithm," ICENCO 2019 2019 15th Int. Comput. Eng. Conf. Util. Mach. Intell. a Better World, pp. 83–86, 2019.

## Business Intelligence for Unemployment Rate Management System

ORIGINA	ALITY REPORT							
SIMILA	8% ARITY INDEX	16% INTERNET SOURCES	8% PUBLICATIONS	4% STUDENT PAPERS				
PRIMAR	Y SOURCES							
1	downloa Internet Source	d.atlantis-press	s.com	9%				
2	Submitted to Forum Komunikasi Perpustakaan Perguruan Tinggi Kristen Indonesia (FKPPTKI) Student Paper							
3	www.de	3%						
4	reposito Internet Source	rium.sdum.umi	nho.pt	3%				

Exclude quotes

On

Exclude matches

< 3%

Exclude bibliography (