# SEMANTIC NETWORK ANALYSIS OF CITIZEN REPORTS FROM INDONESIA'S NATIONAL ONLINE COMPLAINT AND ASPIRATION PORTAL, LAPOR!

**Aliyyah Sarastita Rusdinar[1], Manahan Siallagan[1]**

[1]*School of Business and Management, Institut Teknologi Bandung, Indonesia*
*Email: aliyyah.s@sbm-itb.ac.id*

Abstract

In the era of open government, the Indonesian government with its Open Government Initiative focuses effort on establishing a more participatory democracy. LAPOR! is a national online aspiration and complaint portal set up to put the open government goal into practice. This research aims to provide a structural understanding of one of the practice of open government in Indonesia through a semantic network analysis of the Data gathered from LAPOR!. This research will explore how data gathered from the complaint portal can be (1) used to help identify the on-going issues primarily reported through the LAPOR's Ministry of Law and Human Rights domain, (2) understand the social relationship between the reported issues.

Keywords: Semantic Network Analysis, Open Government, Decision Making, Information and Knowledge Management

## INTRODUCTION

Data has become one of the most vital component of our society's infrastructure. It will continue to be so, as the amount of the digital data increases exponentially overtime. This reality forces the overall conduct of society to gradually transform to accommodate the need of a more open and transparent community. In addition, a world of data influence everyone and everything within it to break out of their silos and live in a more collaborative fashion. The trend for data sharing and interoperability have inevitably outstretched its influence to the government sector and have challenged the very principle of how we run our democracy. As William Eggers described, the shift from industrial age to the information age has propelled governments around the world to transition into adapting a concept of open government (Eggers, 2004). A notion of government practices which embodies four key principles: accountability, transparency, citizen participation and technology and innovation.

This research will specifically highlight how the practice of open government can help promote mutual interaction between government-citizen for a better participatory democracy where more citizen inputs are catered for, thus leading to production of more citizen-centric government policy (Jun & Chung, 2015).

In 2011, Indonesia launched its first step towards the adoption of open government through an initiative popularly known as the Open Government Indonesia, in which the zeal for open government was rendered into several actionable government programs such as Open Data Indonesia, One Map Policy, Open Budget Index, LAPOR!, etc. As to focus on one government program that will best represent the aim of this research, a study will be

conducted to further evaluate the current process of LAPOR!, Indonesia online aspiration and complaint portal.

This study aims to provide a structural understanding of one of the practice of open government in Indonesia through a network analysis of the Data gathered from the online complaint and aspiration portal LAPOR!. This research will explore how data gathered from the complaint portal can be use to help identify the on going issues primarily reported; understand the social relationship between the reported issues; and help map out and recommend priorities in addressing the issues

## LITERATURE REVIEW

### The Challenge of Modern Policy Making

The world is faced with constantly changing realities that offers new sets of problem to be solved. Its nature is more complex as they involve greater aspects coverage, demands urgent resolutions and occur within boundaries of limited world resource. The theory of wicked policy problems, explained how policy making challenges have drastically change as society became more pluralistic, and that a new method of approach is required to address them (Rittel and Webber, 1973). The term "wicked" is used to describe the condition of today policy making challenge that is characterized by the lack of distinct definitions and objectives as the number of influencing variables increase. This circumstance presses for new approach to confront modern policy problems, one that includes intensive consultation among stakeholder in order to synthesize problem framing and sophisticated mathematical analysis to read necessary information (Ferro, Loukis, & Charalabidis, 2013)

Today's policy processes are often scrutinized under the intricate disposition of the constellations in which it belongs to. Various literature sources stress how the challenge of complexity of policy making can be confronted by horizontal policy approach – in which Longo describes as an approach that includes knowledge sharing, collaboration, and the utilization of social networks and computer-supported collaborative policy analysis (Longo, 2013). Although the context of what he discusses is limited to the exertion of policy process within internal formal government, the notion that horizontal approach can and should be initiated as a mind-set to perceive government's external actors. Discerning even individual citizen as pivotal part of the process.

The collaborative nature of the web and the advancement of ICT has enabled government to conduct social listening – the monitoring of ideas and sentiments of citizens engaged in the web – and construct various platform of engagement to directly embrace outside stakeholder in the policy process, albeit mostly only the formulation process. LAPOR! is one good example of Indonesian government attempt to create platform for citizen to reveal on-going problem happening in their life and for government stakeholders to directly communicate with citizen in regards to the problems happening in their units.

However, often times, stakeholder consultation stage produce data that are too large to be manually interpreted. Especially in the case of citizen-government (or generally government-ecosystem stakeholders) interaction, exchanges may result in the production of

big data, data that are random and large in number. For this reason, mathematical analysis needs to be carried to interpret necessary findings related to problem solving. Many research has been conducted to offer possible ways of bridging information and knowledge demands to its practical use to interpret data and assist decision making by making use of information system as a science. This research in particular offers semantic network analysis as one usable tools in processing 'consultation' data from LAPOR! into meaningful information that could directly be used for policy reference.

### *Lapor!*

LAPOR! is a national online aspiration and complaint portal set up to put the goal of increasing citizen engagement in government processes into practice. There are as much as 500 complaints filed on average per day directed to institutions and local government across the country. Reports are inputted through their official website and mobile messaging service and will directly be transferred to relevant government institutions in charge of handling the problems.

Currently, data/report processing and analysis are not the main focus of LAPOR!, however, basic analysis (both manual and automatic) are still carried out. LAPOR!'s internal computerized system processes report statistics and provides visual interpretation of the reports approved. The statistics presents information such as total report input per institutions/work unit, number of report per available word tags; institutions/work unit performance based on the process status (unprocessed, process, solved); etc. Statistics as previously mentioned provide basic information adequate for institutions to monitor work performance of their work units, specific to their responsibility in taking care of citizen complaints. Other than this, LAPOR! current system also provides basic information and attributes of the report filed in form of reports trends, geographic analysis, and categories of report. However, the current in-system statistical depictions are not enough to conclude or show any significant information for policy reference or practice improvement.

### METHODS

### *Data*

Data gathered for the research are historical data in the form of complaints and aspirations imported from LAPOR!'s Ministry of Law and Human Rights disposition tab, spanning from April 2016 to July 2016. Data are reports filed and disposed to Ministry of Law and Human Rights. The ministry is chosen with consideration of their high rate (91%) of case resolution. Institution with high resolution percentage is chosen because it indicates that the institution is active and can best provide real time data of citizen report. In total there are 200 reports data used for analysis.

Due to the variety of data accumulated by LAPOR!, a data cleaning process were manually executed to standardized data. The standardization procedure ensure that data used in the research adheres to the Indonesian formal spelling as is fixed in the KBBI (*Kamus Besar Bahasa Indonesia*, or the Indonesian official dictionary). In overall, this proceeding

will undertake similar process of automatic data normalization that includes case folding, spelling normalization, and filtering process. In the filter process, non-alphabetic symbols along with several type of word (adverbs, pronouns, preposition, articles, conjunction, and particles) were removed. After data are normalized, each report (will be referred to from here on as "cases") are coded into coding group belonging to different categories.

### Semantic Network Analysis

Semantic network analysis is a series of analysis used to map and measure the relationship between words within a measured corpus. Much like social network analysis, semantic network treats words as nodes and co-occurrence ties as social relationship linking those words. By observing the relational construct between words, this research attempts to unearth the trends of issues among the report data used and their centrality in the overall networks. The statistical techniques used will cover: the term frequency-inverse document frequency, construction semantic matrix of similarity and semantic network, hierarchical clustering, and degree centrality measurement.

### Term frequency-inverse document frequency (TF.IDF)

TF.IDF measures the importance of a term relative to a document or a collection of documents. The basic understanding of term frequency is when within a document or any measured textual data scope a query term is mentioned numerous times, the likelihood of the term to have more relevance to the measured textual data scope. However, measuring relevance by term frequency alone would be feeble if one does not consider its inverse document frequency. A collection of documents containing the topic of human anatomy almost always will have the word "human" in each document. Measuring term frequency alone would be insufficient as the TF measurement treats all term as equally important. Inverse document frequency serves as a mechanism to debilitate the effect of a query term occurring too much in the measured collection of documents.

TF.DF (Term Frequency Inverse - Document Frequency) is measured:

$$tf.idf_{t,d} = tf_{t.d} \; x \; idf_t$$

### Similarity Measures

Similarity statistics illustrate the relationship between words by examining the count of ties a node (in this case, a keyword) has that are similar to those of other nodes (in some types of similarity matrices, the values within the cells are represented in percentage). In the case where network is scattered, it would be quite a predicament to distinguish structural equivalence. The Jaccard coefficient approach is to calculate the number of times that both nodes record a tie to another node (a third note).

Similarity measures recorded in a symmetrical matrix of similarity will be used for hierarchical clustering and semantic network analysis.

### *Hierarchical Clustering (Dendogram)*

Hierarchical clustering will be presented with a dendogram, a diagram that represents similarity between entities. Dendogram can be used to measure similarity between any entity. In this case, this research takes a lexomic approach and will try to measure the similarity between words (with the parameter of occurrence) to further explain the relationship between the coding and chosen keyword complaints or aspirations reports that has been chosen based on their relevance (measured with TD.IDF) by grouping them into clusters of issue.

### *Degree Centrality Measurement*

Degree calculation measures centrality by the amount of connection a word has with other words. A word is seen to have a central presence within a network when its connectivity with other words in network is high. By measuring degree centrality, this research will try to understand which issues is considered most central.

## DATA ANALYSIS

Data analysis will be divided into two main segment which will cover the analysis of Ministry of Law and Human Rights' cases coding (overall) and cases belonging to The Department of Immigration category.

### *Code Analysis*
### Coded Data Distribution

Table 1. Coded Case Distribution

| Category | Count | %Count | Cases | %Cases |
|---|---|---|---|---|
| Department of Immigration | 194 | 76.08% | 186 | 75.30% |
| Ministry of Law and Human Rights (general) | 11 | 4.31% | 11 | 4.45% |
| Department of General Law Administration | 15 | 5.88% | 15 | 6.07% |
| Department of Civic Affairs | 28 | 10.98% | 28 | 11.34% |
| Ministry of Law and Human Rights Branch Office | 3 | 1.18% | 3 | 1.21% |
| Institution of National Law Assistance (BPHN) | 1 | 0.39% | 1 | 0.40% |
| Department of Intellectual Property Rights | 2 | 0.78% | 2 | 0.81% |
| Special Cases Ministry of Law and Human Rights | 1 | 0.39% | 1 | 0.40% |
| Total | 255 | 100.00% | 247 | 100.00% |

As shown in table 1, there are 8 categories of code and within each of them, contains different number of code groups. All 200 cases were coded with at least one code per case. There are cases that contains more than one code and segments within cases sometimes are coded more

than once into different coding groups. The table reveals that 75.3% of the cases are coded as "Department of Immigration", while the remaining 24.7% are distributed within the other 7 code group. Based on this findings, this research will proceed later in the next analysis segment, with only processing to analyse data that are coded as the "Department of Immigration" as the other coded group are deemed as having too little data to be properly analyse.

### *Hierarchical Clustering of Code*
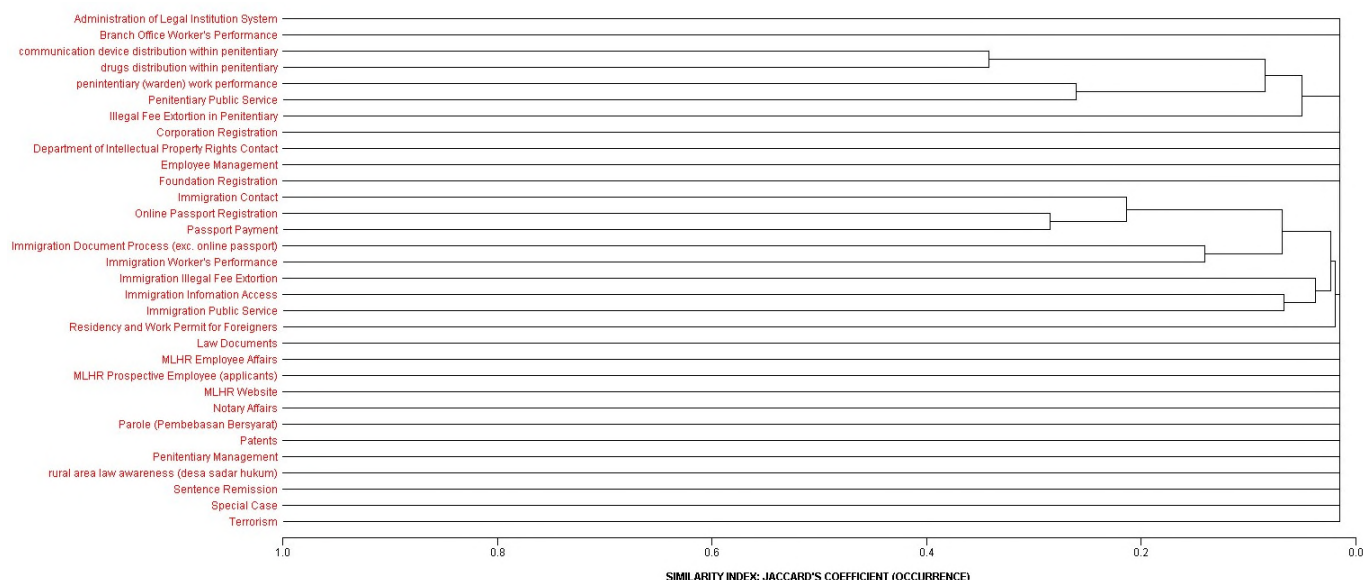


Figure 1. Dendogram of Code

The figure 1 shows the hierarchical clustering of the case coding, represented in dendogram. Clustering is done by examining coding co-occurrence within cases using Jaccard's coefficient. The dendogram reveals that out of the 31 codes, there are 19 codes whose distribution within cases are substantially different and thus, does not intercede with each other. Among these are cases coded with: Terrorism, Special Case, Rural Area Law Awareness, Penitentiary Management, Patents, Parole, Notary Affairs, Ministry of Law and Human Rights (MLHR) Website, MLHR Prospective Employee, MLHR employee affairs, Law Documents, Foundation Registration, Employee Management, Department of Intellectual Property Rights, Corporation Registration, Branch Office's Workers Performance, and Administration of Legal Documents System. Although, some of the mentioned code are listed within the same category they do not seem to have ever been coded within the same case, indicating absence of potential relation. However, it is important to note that within some of the coding group (refer to Table 1A Appendix) only consist of very little amount of data. This render any effort to make conclusion based on this inadequate. More can potentially reveal more relationship between coding group of the same category.

In the other hand, the dendogram also reveals two main clades that elucidate relationship between the other 12 codes. On brief observation, it is reveal that the two clades illuminate coding relationship between two separate categories, that is codes within

Department of Immigration and Department of Civic Affairs. Within the Department of Civic Affairs branch, it is seen that Drugs Distribution Within Penitentiary is most similar to Communication Device Distribution Within Penitentiary (Cluster I), and although their distance is quite great, they appear to be the pair with the shortest distance relative to other pairs of code. This shows that the two codes the ones that are mostly likely to appear together within a case. The next pair of code arranged next to the cluster I are Warden's Work Performance and Penitentiary Public Service (cluster II) displaying similar relationship as cluster I but with a slightly higher degree of dissimilarity. Cluster I and II formed another clade that reveals that Drugs Distribution Within Penitentiary and Communication Device Distribution Within Penitentiary along with Warden's Work Performance and Penitentiary Public Service share a similarity relationship. A conclusion can be made that if coding represents issues, issue relating distribution of drugs and communication device within prison cells may well be related to issue regarding prison's warden work performance (be it negligence or voluntary cooperation) and penitentiary public service. Moving on from that, it is depicted that cluster III (a clade branching out to cluster I and II) connects to the code Illegal Fee Extortion (cluster IV), although the similarity degree is very low. This shows that even though cluster III and IV is somehow similar, the distribution of Illegal Fee Extortion is essentially different. In spite of the fact that the result may have appear to be this way due to the lack of cases measured within the Illegal Fee Extortion code.

The second main clade illustrated is one that likely has to do with Department of Immigration category. It is revealed that Online Passport Registration and Passport Payment (cluster V) are has the shortest distance of similarity, and is also part of a bigger cluster along with Immigration Contact (cluster VI). From this finding, it could be said that a report filed regarding online passport registration also reports problem regarding passport payment, and that reports containing both issues will also likely to contain problem relating to immigration contact. Moving upwards, cluster VI seems to be a part of another bigger clave of cluster VIII, where it is grouped with cluster VII. A clade that branches to the code Immigration Document Process (excluding online passport) and Immigration's Worker's performance. Their distance of similarity is quite far, so the conclusion that both issues are likely to be reported together is feeble. The similarity that cluster VI and VII share is also depicted with a very long branch, indicating that it would be inappropriate to come to any conclusion expect that the two cluster are somehow similar to each other. Other cluster similarity relationship (cluster, VIII, IX, X and XI moving top to bottom, left to right) are almost insignificant as their distance is extremely long. However, they are all revealed to be part of the same category, which is the Department of Immigration.

### *Department of Immigration's Cases Analysis*

As has been mentioned previously, this research will conduct a categorical data analysis for the category with the most significant portion of cases. Coded data distribution table shows that 75.4% of the cases are coded into Department of Immigration category. Therefore, this section will elaborate the findings discovered when keywords of cases within the Department

of Immigration category are used as a measured corpus (collection of cases). There is a total of 5,655 words with 941 word forms. However, to simplify the analysis process this research will look at word with different re-occurring frequency, meaning only some percentage of the total words will be relevant for further discussion depending on stages of the analysis. The cut is decided by simply iterating the word cutting process until a dominant pattern can be interpreted.

### Term Frequency – Inverse Document Frequency (TF.IDF)

LAKU, BAYAR, PASPOR, IMIGRASI, MOHON, PROSES, DATANG, KANIM, ONLINE and BANK appears to be the top ten most relevant words within the Department of Immigration category (refer to table 2A in Appendix). Ranging from 67.8 as the highest score to 46.9 as the lowest score out of all the ten keywords. The finding of relevant keyword would serve as a primary datum to further interpret possible issues. A more comprehensive interpretation can only be drawn when the relationship (similarity or distance) between these high-scoring keywords with each other or with keywords of lesser relevance is revealed.
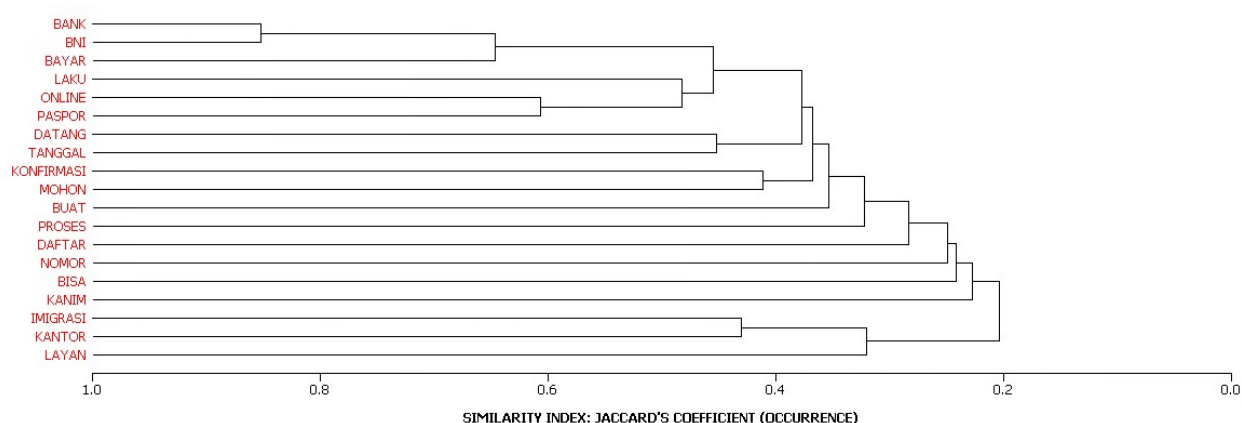


Figure 5 Department of Immigration Dendogram

Figure 2 displays that BANK and BNI ranks first as the pair of word with highest degree of similarity. The word BANK is highly likely to occur along with the word BNI within a case, as their degree of similarity is high, 0.854, BANK BNI (Cluster 1) will be referred together from here on as a cluster of their own to signify a new and separate concept, as opposed to valuing them with individual meaning (BANK and BNI). The cluster BANK BNI (Cluster I) is apparently shown to share a close similarity with the word BAYAR, with degree of similarity of 0.644, forming a new cluster of BAYAR BANK BNI (Cluster II). The formation of this particular cluster has generated BAYAR BANK BNI (loosely translated to Bank BNI Payment) as the first observable issue within Department of Immigration. Among the three words of the first determined issue, only one of the word (BNI) are not part of the top ten highest scoring TD.IDF calculation, this helps draw the assumption that the absence of the word BNI would not significantly alter the substance of

the determined issues. BAYAR BANK BNI and BAYAR BANK would more or less imply to similar context.

Right below Cluster II, it is displayed that the pair with the next closest degree of similarity is ONLINE and PASPOR, forming a cluster of ONLINE PASPOR (Cluster III) with degree value of 0.604. Cluster III is then followed by the formation of Cluster IV, where with degree value of 0.477 it unites with the word LAKU, making a LAKU ONLINE PASPOR cluster. All three words are within the top ten high scoring TD.IDF list, making the relevance of all word almost equally significant. This cluster, LAKU ONLINE PASPOR (literally translated to online passport perform or loosely interpreted as online passport) can also be determined as an observable issue it its own.

Figure 2 continues to display the relationship between the two determined issue Payment Bank BNI (1) and Online Passport (2). Cluster II and Cluster IV forming Cluster V indicates that report about online passport has 44.9% likelihood of occurring together with reports about Bank BNI payment, almost half of the time. In practice, payment at Bank BNI is one of the procedure of online passport application, this finding shows that within the process of online passport application payment process is one of the frequently reported issues (keep in mind that this research does not inquire the sentiment of the report and merely dismantle the extend of similarity and likelihood of co-occurrence of the reported issues).

The rest of the Dendogram further reveals the relationship between other words, so far, it can be seen that DATANG is paired with TANGGAL (Cluster VI) with degree similarity value of 0.446, MOHON with KONFIRMASI (Cluster VII) with degree similarity value of 0.404, and IMIGRASI with KANTOR (Cluster VIII) with degree similarity value of 0.42. Among these clusters, DATANG TANGGAL seems to be the pair arranged closest to Cluster V (Cluster IX). Indicating that DATANG TANGGAL (loosely translated and interpreted as visitation date (3)), another determinable issues shares 37% likelihood of being reported together in reports in which issues of online passport and Bank BNI payment are co-occurring. Apparently, in practice visitation to immigration office for interview and photoshoot session is also a part of passport application process, and the visitation date is sent to applicant through email only after payment via bank transfer has been confirmed. This fact is enforced by the finding that MOHON KONFIRMASI (loosely translated into confirmation plea (4)) comes next in the arrangement, displaying 36% of similarity with Cluster IX (Cluster X), determining another observable issue reported. In summary, Cluster X can be said to contain mostly issues regarding general passport application procedure.

The next few branches do not seem to lead to potential issue revelation as a degree of similarity and therefore will not be included in the analysis. In the lowest part of the diagram, another Cluster VII (KANTOR IMIGRASI, loosely translated and interpreted as Immigration Office) is seen to have connect itself with the word LAYAN (service). As the combination of words produce in this cluster create a new context when used together (Immigration Office Service (5)), this will be determined as another observable issues. However, table 3A (Appendix) shows that there is only 19.3% likelihood that report containing immigration office service will occur together with report of passport application

procedure which also at the same time discuss online passport, bank payment, visitation date, and confirmation plea. This makes sense as have been hinted by the Dendogram arrangement, IMIGRASI KANTOR LAYAN are located in the furthest bottom of the diagram, signifying their substantially different distribution within cases.

### *Semantic Network of Similarity*

Figure 1A (Appendix) shows that BAYAR BANK BNI and LAKU PASPOR ONLINE are example of clusters connection that are not only highly similar in terms of their likelihood of co-occurring together but also in their connectivity pattern with other words. The fact that the two cluster is positioned relatively central may indicate that they are indeed central issues within the network. To confirms this, this research continues to evaluate each word's degree and eigenvector centrality which confirms the most central issue reported within the Department of Immigration is Payment Bank BNI (represented by the words BAYAR BANK BNI), followed by Online Passport, Visitation date and Plea for Confirmation (refer to Figure 2A, Appendix for centrality value)

## CONCLUSION

### *Case distribution and issue trends within Ministry of Law and Human Rights*

There are two noteworthy discoveries uncovered from the code analysis. Firstly, it is found that the distribution of reports directed to Ministry of Law and Human Rights from LAPOR! is concentrated within the domain of work of Department of Immigration (75.3%). Based on the hierarchical clustering analysis of cases coding, relatively significant degrees of similarity of code co-occurrence between the code: online passport registration, passport payment, immigration contact were discovered. It shows that reports filed to Ministry of Law and Human Rights regarding online passport registration are also likely to report issue regarding passport payment 27% of the time. While reports containing both issues will also 20.2% likely to contain problem relating to immigration contact.

Secondly, although with significant difference, the next densely populated codes are within the domain of work of Department of Civic Affairs (11.34%). It is uncovered that among codes within Department of Civic Affairs, visible degree of similarity of code co-occurrence are found between the code: drugs distribution within penitentiary, communication device distribution within penitentiary, warden's work performance, and penitentiary public service. It is interpreted that reports filed about drugs distribution within penitentiary are also 33.3% likely to mention about communication device distribution within penitentiary. While reports containing both issues will 7.71% likely to report warden's work performance and penitentiary public service.

Although code analysis is able to expose the concentration pattern of cases within Ministry of Law and Human Rights and the relationship between some of the codes, it was not able to give representative statistics that allows conclusion of trends of issue to be drawn directly. This is due to the fact that not enough cases were analysed. The analysis will

produce better result by showing greater degree of issue variety and a more precise similarity relationship measurement, if the number of cases measured are increased.

### *Trend of Issues and Structural Construct of Issue within Department of Immigration*
The category with the most dominant case distribution was used for further analysis in order to reveal a more statistically precise result in determining trends of issue reported to Ministry of Law and Human Rights. Department of Immigration was chosen as it displays the highest percentage of case concentration. The hierarchical clustering analysis conducted reveals that there are 5 observed issues within Department of Immigration which varies in level of significance, the issues are: (1) Bank BNI Payment, (2) Online Passport, (3) Visitation Date, (4) Plea for Confirmation, (5) Immigration Office Service. Using the visualization of similarity network and the calculation of degree centrality to explore the structural construct of the similarity network, this research is able to conclude that the most central issue reported within the Department of Immigration is Online Passport and Payment Bank BNI.

## REFERENCES

Chae Nam Jung, C. J. (2005). Big Data Analysis of Local Government 3.0: Focusing on Gyeongsangbuk-do in Korea. *Technological Forecasting and Social Change*.

Eggers, W. (2004). *Government 2.0: Using Technology to Improve Education, Cut Red Tape, Reduce Gridlock, and Enhance Democracy*. Rowman & Littelfield.

Ferro, E., Loukis, E., & Charalabidis, Y. (2013). Policy Making 2.0: From Theory to Practice. *Government Information Quarterly*.

Hanneman, R. A. (2005). *Introduction to Social Network* .

Longo, J. (2013). *Towards Policy Analysis 2.0.*

Rittel, H. W. (1973). Dillemas in General Theory of Planning. *Policy Sciences*, 155-169.

## APPENDIX

Table 1A Full Category and Code List

| Category | Code | N | % Codes | Cases | % Cases |
|---|---|---|---|---|---|
| Department of Immigration | Online Passport Registration | 65 | 25.70% | 62 | 31.00% |
| Department of Immigration | Immigration Worker's Performance | 9 | 3.60% | 9 | 4.50% |
| Department of Immigration | Immigration Contact | 22 | 8.70% | 21 | 10.50% |
| Department of Immigration | Immigration Document Process (exc. online passport) | 36 | 14.20% | 34 | 17.00% |
| Department of Immigration | Passport Payment | 27 | 10.70% | 26 | 13.00% |
| Department of Immigration | Immigration Infomation Access | 1 | 0.40% | 1 | 0.50% |
| Department of Immigration | Residency and Work Permit for Foreigners | 8 | 3.20% | 8 | 4.00% |
| Department of Immigration | Immigration Public Service | 19 | 7.50% | 18 | 9.00% |
| Department of Immigration | Immigration Illegal Fee Extortion | 5 | 2.00% | 5 | 2.50% |
| Ministry of Law and Human Rights (general) | MLHR Website | 6 | 2.40% | 6 | 3.00% |
| Ministry of Law and Human Rights (general) | MLHR Prospective Employee (applicants) | 1 | 0.40% | 1 | 0.50% |
| Ministry of Law and Human Rights (general) | Terrorism | 1 | 0.40% | 1 | 0.50% |
| Ministry of Law and Human Rights (general) | MLHR Employee Affairs | 2 | 0.80% | 2 | 1.00% |
| Department of General Law Administration | Foundation Registration | 6 | 2.40% | 6 | 3.00% |
| Department of General Law Administration | Administration of Legal Institution System | 1 | 0.40% | 1 | 0.50% |
| Department of General Law Administration | Notary Affairs | 3 | 1.20% | 3 | 1.50% |
| Department of General Law Administration | Corporation Registration | 2 | 0.80% | 2 | 1.00% |
| Department of General Law Administration | Law Documents | 3 | 1.20% | 3 | 1.50% |
| Department of Civic Affairs | drugs distribution within penitentiary | 4 | 1.60% | 4 | 2.00% |
| Department of Civic Affairs | penintentiary (warden) work performance | 5 | 2.00% | 4 | 2.00% |
| Department of Civic Affairs | communication device distribution within penitentiary | 4 | 1.60% | 4 | 2.00% |
| Department of Civic Affairs | Sentence Remission | 4 | 1.60% | 4 | 2.00% |
| Department of Civic Affairs | Penitentiary Public Service | 1 | 0.40% | 1 | 0.50% |
| Department of Civic Affairs | Penitentiary Management | 5 | 2.00% | 5 | 2.50% |
| Department of Civic Affairs | Parole (Pembebasan Bersyarat) | 2 | 0.80% | 2 | 1.00% |
| Department of Civic Affairs | Illegal Fee Extortion in Penitentiary | 4 | 1.60% | 4 | 2.00% |
| Ministry of Law and Human Rights Branch Office | Branch Office Worker's Performance | 2 | 0.80% | 2 | 1.00% |
| Ministry of Law and Human Rights Branch Office | Employee Management | 1 | 0.40% | 1 | 0.50% |
| Institution of National Law Assistance (BPHN) | rural area law awareness (desa sadar hukum) | 1 | 0.40% | 1 | 0.50% |
| Department of Intellectual Property Rights | Department of Intellectual Property Rights Contact | 1 | 0.40% | 1 | 0.50% |
| Department of Intellectual Property Rights | Patents | 1 | 0.40% | 1 | 0.50% |
| Special Cases Ministry of Law and Human Rights | Special Case | 1 | 0.40% | 1 | 0.50% |

Table 2A TF.IDF Table (Words with Min. Frequency 45)

| WORD | FREQUENCY | % TOTAL | NO. CASES | % CASES | TF • IDF |
|------|-----------|---------|-----------|---------|----------|
| *LAKU* | 139 | 2.46% | 65 | 32.50% | 67.8 |
| *BAYAR* | 127 | 2.25% | 59 | 29.50% | 67.3 |
| *PASPOR* | 219 | 3.87% | 103 | 51.50% | 63.1 |
| *IMIGRASI* | 119 | 2.10% | 65 | 32.50% | 58.1 |
| *MOHON* | 88 | 1.56% | 45 | 22.50% | 57 |
| *PROSES* | 74 | 1.31% | 39 | 19.50% | 52.5 |
| *DATANG* | 81 | 1.43% | 48 | 24.00% | 50.2 |
| *KANIM* | 68 | 1.20% | 37 | 18.50% | 49.8 |
| ONLINE | 103 | 1.82% | 67 | 33.50% | 48.9 |
| *BANK* | 65 | 1.15% | 38 | 19.00% | 46.9 |
| *BUAT* | 80 | 1.41% | 52 | 26.00% | 46.8 |
| *KANTOR* | 54 | 0.95% | 29 | 14.50% | 45.3 |
| *NOMOR* | 52 | 0.92% | 28 | 14.00% | 44.4 |
| *DAFTAR* | 62 | 1.10% | 39 | 19.50% | 44 |
| *KONFIRMASI* | 45 | 0.80% | 28 | 14.00% | 38.4 |
| *LAYAN* | 45 | 0.80% | 29 | 14.50% | 37.7 |
| *TANGGAL* | 46 | 0.81% | 33 | 16.50% | 36 |
| *BISA* | 48 | 0.85% | 37 | 18.50% | 35.2 |
| *BNI* | 48 | 0.85% | 38 | 19.00% | 34.6 |

Table 3A Similarity Value of Code Dendogram

| CLUSTER | GROUP 1 | GROUP 2 | SIM |
|---------|---------|---------|-----|
| 1 | communication device distribution within penitentiary | drugs distribution within penitentiary | 0.333 |
| 2 | Online Passport Registration | Passport Payment | 0.275 |
| 3 | penitentiary (warden) work performance | Penitentiary Public Service | 0.25 |
| 4 | Immigration Contact | Cluster 2 | 0.202 |
| 5 | Immigration Document Process (non online) | Immigration Worker's Performance | 0.128 |
| 6 | Cluster 1 | Cluster 3 | 0.071 |
| 7 | Cluster 4 | Cluster 5 | 0.054 |
| 8 | Immigration Information Access | Immigration Public Service | 0.053 |
| 9 | Cluster 6 | Illegal Fee Extortion in Penitentiary | 0.036 |
| 10 | Immigration Illegal Fee Extortion | Cluster 8 | 0.023 |
| 11 | Cluster 7 | Cluster 10 | 0.008 |
| 12 | Cluster 11 | Residency/Work Permit for Foreigners | 0.004 |
| 13 | Special Case | Terrorism | 0 |
| 14 | Sentence Remission | Cluster 13 | 0 |
| 15 | rural area law awareness (Desa Sadar Hukum) | Cluster 14 | 0 |
| 16 | Penitentiary Management | Cluster 15 | 0 |
| 17 | Patents | Cluster 16 | 0 |
| 18 | Parole (Pembebasan Bersyarat) | Cluster 17 | 0 |
| 19 | Notary Affairs | Cluster 18 | 0 |
| 20 | MLHR Website | Cluster 19 | 0 |
| 21 | MLHR Prospective Employee (applicants) | Cluster 20 | 0 |
| 22 | MLHR Employee Affairs | Cluster 21 | 0 |
| 23 | Law Documents | Cluster 22 | 0 |
| 24 | Cluster 12 | Cluster 23 | 0 |
| 25 | Foundation Registration | Cluster 24 | 0 |
| 26 | Employee Management | Cluster 25 | 0 |
| 27 | Department of Intellectual Property Rights Contact | Cluster 26 | 0 |
| 28 | Corporation Registration | Cluster 27 | 0 |
| 29 | Cluster 9 | Cluster 28 | 0 |
| 30 | Branch Office Worker's Performance | Cluster 29 | 0 |
| 31 | Administration of Legal Institution System | Cluster 30 | 0 |

Table 4A Department of Immigration Cases Dendogram

| Cluster | Group 1 | Group 2 | Similarity |
|---------|---------|---------|-----------|
| I | BANK | BNI | 0.854 |
| II | Cluster I | BAYAR | 0.644 |
| III | ONLINE | PASPOR | 0.604 |
| IV | LAKU | Cluster III | 0.477 |
| V | Cluster II | Cluster IV | 0.449 |
| VI | DATANG | TANGGAL | 0.446 |
| VII | IMIGRASI | KANTOR | 0.424 |
| VIII | KONFIRMASI | MOHON | 0.404 |
| IX | Cluster V | Cluster VI | 0.37 |
| X | Cluster IX | Cluster VIII | 0.36 |
| XI | Cluster X | BUAT | 0.345 |
| XII | Cluster XI | PROSES | 0.314 |
| XIII | Cluster VII | LAYAN | 0.312 |
| XIV | Cluster XII | DAFTAR | 0.274 |
| XV | Cluster XIV | NOMOR | 0.239 |
| XVI | Cluster XV | BISA | 0.231 |
| XVII | Cluster XVI | KANIM | 0.217 |
| XVII | Cluster XVII | Cluster XIII | 0.193 |
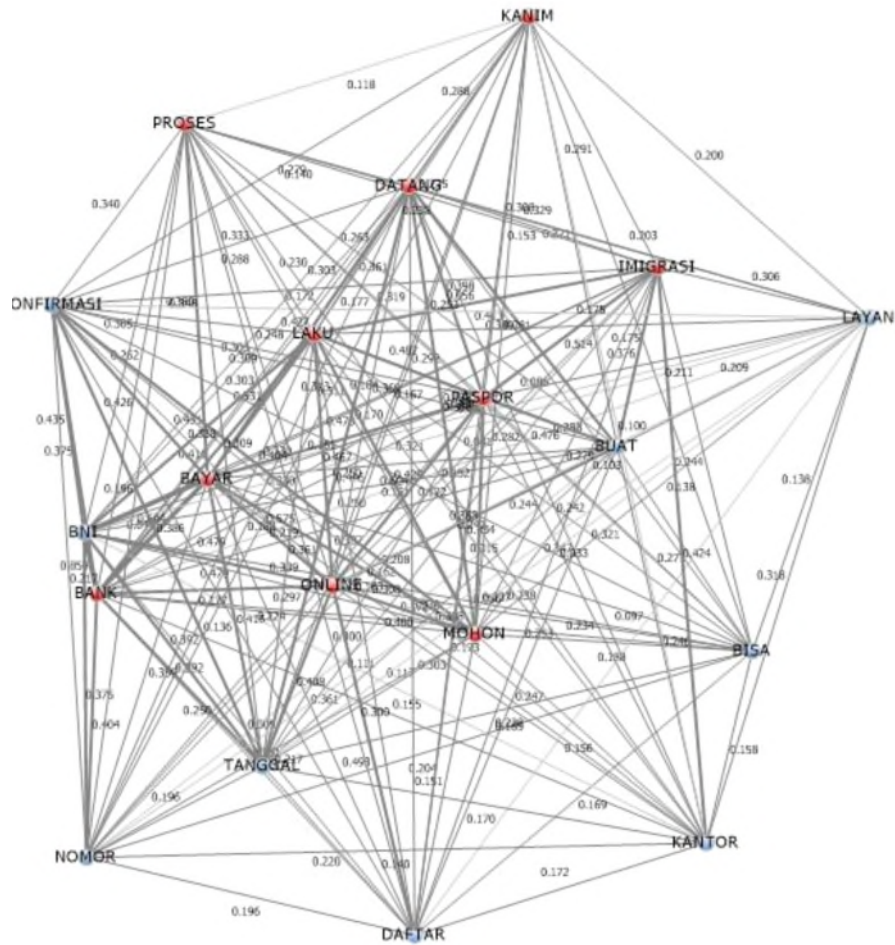
Figure 1A Semantic Network of Similarity Department of Immigration

|   |   | 1<br>Degree | 2<br>nDegree |
|---|---|---|---|
|   |   | ----- | ----- |
| 1 | BANK | 6.160 | 0.401 |
| 2 | BAYAR | 7.072 | 0.460 |
| 3 | BISA | 3.947 | 0.257 |
| 4 | BNI | 6.226 | 0.405 |
| 5 | BUAT | 5.496 | 0.358 |
| 6 | DAFTAR | 4.318 | 0.281 |
| 7 | DATANG | 5.697 | 0.371 |
| 8 | IMIGRASI | 5.160 | 0.336 |
| 9 | KANIM | 3.845 | 0.250 |
| 10 | KANTOR | 3.562 | 0.232 |
| 11 | KONFIRMASI | 4.949 | 0.322 |
| 12 | LAKU | 6.340 | 0.412 |
| 13 | LAYAN | 2.644 | 0.172 |
| 14 | MOHON | 5.397 | 0.351 |
| 15 | NOMOR | 3.833 | 0.249 |
| 16 | ONLINE | 6.681 | 0.435 |
| 17 | PASPOR | 6.455 | 0.420 |
| 18 | PROSES | 4.736 | 0.308 |
| 19 | TANGGAL | 5.280 | 0.343 |

```
Graph Centralization -- a

                1
             Centra
             lizati
               on
             ------
    1 Sheet1 0.1399
```
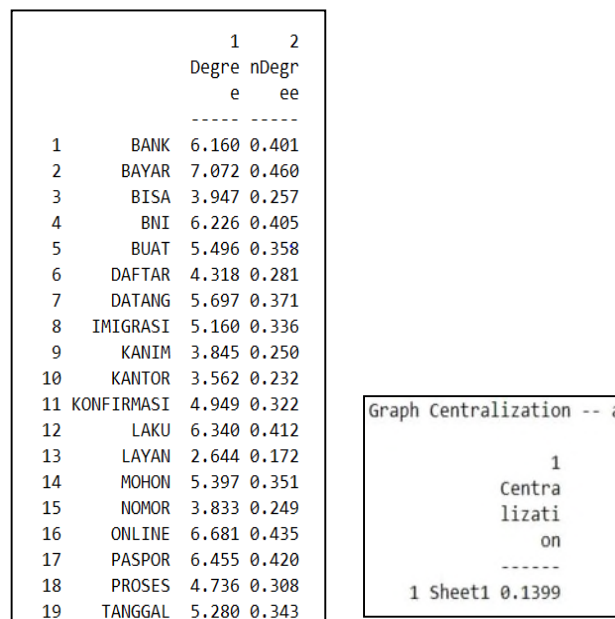
Figure 2A Degree Centrality Measure Department of Immigration