

**PENERAPAN METODE *ONTOLOGY ENRICHMENT* DENGAN
KONSEP AMENITAS PADA ONTOLOGI DI DOMAIN *TOURISM*
(*case study: DWIPA Ontology*)**

TUGAS AKHIR



SHANIA ISYHRANI

1142001008

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS BAKRIE
JAKARTA**

2019

**PENERAPAN METODE *ONTOLOGY ENRICHMENT* DENGAN
KONSEP AMENITAS PADA ONTOLOGI DI DOMAIN *TOURISM*
(*case study: DWIPA Ontology*)**

TUGAS AKHIR

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer**



SHANIA ISYHRANI

1142001008

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN ILMU KOMPUTER
UNIVERSITAS BAKRIE
JAKARTA
2019**

HALAMAN PERNYATAAN ORISINILITAS

Tugas akhir ini adalah benar hasil karya sendiri,
dan semua sumber baik yang dikutip maupun dirujuk,
telah saya nyatakan dengan benar.

Nama : Shania Isyahrani

NIM : 1142001008

Tanda Tangan : 

Tanggal : 21 Agustus 2019

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama : Shania Isyahrani
NIM : 1142001008
Program Studi : Informatika
Fakultas : Teknik dan Ilmu Komputer
Judul Tugas Akhir : Penerapan Metode *Ontology Enrichment* dengan konsep Amenitas pada Ontologi di *Domain Tourism (case study: DWIPA Ontology)*

Telah berhasil dipertahankan dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika Fakultas Teknik dan Ilmu Komputer, Universitas Bakrie.

DEWAN PENGUJI

Pembimbing I : Guson P. Kuntarto, S.T., M. Sc.
Pembimbing II : Irwan P. Gunawan, S.T., M.Eng, PhD
Penguji : Ir. Kenny Badjora Lubis, M.Kom
Penguji : Siti Rohajawati, S.Kom, M.Kom, Dr.


(.....) 27/8

(.....) 27/8

(.....)

(.....)

Ditetapkan di : Jakarta

Tanggal : 21 Agustus 2019

UNGKAPAN TERIMA KASIH

Alhamdulillahirabbil'aalamiin, puji dan syukur kehadiran Allah SWT atas segala rahmat dan karunia-Nya Tugas Akhir yang berjudul "Penerapan Metode *Ontology Enrichment* dengan konsep Amenitas pada Ontologi di Domain *Tourism (case study: DWIPA Ontology)*" dapat diselesaikan. Shalawat dan salam penulis haturkan kepada Nabi Besar Muhammad SAW, keluarga dan para sahabatnya yang telah membimbing umatnya ke masa yang terang benderang penuh dengan cahaya iman.

Penyusunan Tugas Akhir ini tidak lepas dari berbagai hambatan dan kesulitan dari awal hingga akhir penyusunan. Penulis menyampaikan terima kasih yang sebesar-besarnya kepada Kedua Orangtua yang terhormat dan tercinta, Bapak Syaiful Alamsyah dan almarhumah Ibu Lilia Iswahyuni, yang selalu memberikan dukungan, motivasi dan doa yang senantiasa mengiringi langkah penulis, serta Saudara, Leo Isyahputra yang mendukung dan memberikan semangat kepada penulis. Begitu banyak pihak yang turut membantu dan memberikan dukungan, masukan, nasihat, serta doa selama penyusunan Tugas Akhir ini. Oleh karena itu, dengan segala hormat, penulis mengungkapkan rasa terima kasih kepada:

1. Bapak Guson P. Kuntarto, S.T., M. Sc. selaku dosen Pembimbing 1 yang telah mengerahkan ide dan gagasan untuk membantu proses pengerjaan dan penyusunan Tugas Akhir.
2. Bapak Irwan P. Gunawan, S.T., M.Eng, PhD selaku dosen Pembimbing 2 Tugas Akhir yang telah meluangkan waktu serta memberikan saran untuk membantu proses pengerjaan dan penyusunan Tugas Akhir.
3. Bapak Berkah I. Santoso dan Bapak Yusuf Lestanto selaku dosen Pembahas Seminar Tugas Akhir, yang memberikan motivasi, saran, serta perbaikan terhadap penyusunan Tugas Akhir.

4. Bapak Ir. Kenny Badjora Lubis, M.Kom. dan Ibu Siti Rohajawati, S.Kom, M.Kom, Dr. selaku dosen Penguji Sidang Tugas Akhir yang memberikan saran dan perbaikan terhadap Tugas Akhir.
5. Yossy Alrin dan Muh Arfandy Wijaya AK yang selalu ada dan mendukung, membantu serta memberikan motivasi kepada penulis selama pengerjaan Tugas Akhir ini dari awal sampai saat ini.
6. Teman-teman Informatika 2014, Putri Nurrahmah Arta yang selalu membantu dan mendukung penulis dalam suka dan duka, Rizky Aeamelia, Maissy, Safira Alisha Aqista, dan Rizka yang selalu memberikan motivasi dan dukungan, serta Arisyi Falih Naufal, Imam Sopani, Muhammad Rahman Muttaqin dan Muhammad Fauzan Adhim yang selalu memberikan semangat kepada penulis.
7. Kakak-kakak Informatika Universitas Bakrie Angkatan 2013 yang telah membagikan pengalaman, memberikan saran dan motivasi serta semangat selama penyusunan Tugas Akhir.
8. Adik-adik Informatika Universitas Bakrie Angkatan 2015 yang telah memberikan doa, dukungandan semangat kepada penulis selama penyusunan Tugas Akhir.
9. Seluruh pihak Program Studi Informatika Universitas Bakrie yang telah memberikan ilmu dan pembelajaran serta pengalaman yang sangat bermanfaat bagi penulis selama masa perkuliahan.

Semoga Allah SWT senantiasa membalas kebaikan dan memberikan keberkahan kepada kita semua. Penulis berharap Tugas Akhir ini dapat memberi informasi yang berguna dan bermanfaat bagi berbagai kalangan bidang pendidikan, khususnya bidang informatika.

Jakarta, 21 Agustus 2019

Shania Isyahrani

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai *civitas* akademik Universitas Bakrie, saya yang bertanda tangan di bawah ini:

Nama : Shania Isyahrani
NIM : 1142001008
Program Studi : Informatika
Fakultas : Teknik dan Ilmu Komputer
Jenis Tugas Akhir : Penerapan Metode

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bakrie **Hak Bebas Royalti Noneksklusif** (*Non-exclusive Royalty-Free Right*) atas tugas akhir saya yang berjudul:

**Penerapan Metode *Ontology Enrichment* dengan konsep Amenitas pada
Ontologi di Domain *Tourism*
(*case study: DWIPA Ontology*)**

Dengan Hak Bebas Royalti Noneksklusif ini, Universitas Bakrie berhak menyimpan, mengalihmedia/formatkan, merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta untuk kepentingan akademis.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 21 Agustus 2019



Shania Isyahrani

**PENERAPAN METODE *ONTOLOGY ENRICHMENT* DENGAN KONSEP
AMENITAS PADA ONTOLOGI DI DOMAIN *TOURISM***
(case study: *DWIPA Ontology*)

Shania Isyahrani

ABSTRAK

Menanggapi perkembangan *Tourism* di Indonesia, Ontologi DWIPA III dibangun untuk merepresentasikan domain *tourism* di Indonesia khususnya pada daerah Bali. Ontologi DWIPA III terdiri dari *Class/subClass Accommodations, Attractions, Regencies* dan *Events*. *Tourism* sendiri memiliki 3 elemen penting dalam perkembangannya, yaitu *Attractions, Accommodations* dan *Amenities*. Terkait dengan hal tersebut, Ontologi DWIPA III masih memiliki kekurangan pada konsep *Amenities* yang mendukung *tourism* itu sendiri. Penelitian ini menggunakan metode *Ontology Enrichment* untuk menambahkan konsep baru kedalam ontologi DWIPA III. Sumber data amenitas diambil dari situs *tourism* yang menyediakan informasi amenitas yaitu TripAdvisor. Data amenitas yang telah dikumpulkan lalu diberi fitur untuk membedakan data amenitas dan bukan amenitas, yang selanjutnya dilakukan pembobotan fitur menggunakan metode *Terms Frequency – Inverse Document Frequency*. Selanjutnya data akan di klasifikasi dengan membandingkan algoritma *K-Means, K-Nearest Neighbor* dan *Self-Organizing Maps*. Nilai *Kappa Value* yang didapatkan dari ketiga algoritma yang digunakan menunjukkan algoritma *K-Means* memiliki keakuratan sebesar 0.69 dinyatakan sebagai *Substantial Agreement*. Selanjutnya Ontologi akan dievaluasi menggunakan *Data Driven* dan *OntoQA*. Hasil akhir penelitian ini menunjukkan bahwa Ontologi DWIPA III berhasil diperkaya dengan memiliki 4 *main Class*, 29 *subClass* dan 319 *instance*, dengan nilai *Relationship Richness* sebesar 0.33, *Inheritance Richness* sebesar 32 dan *Attribute Richness* sebesar 79.

Kata Kunci: Ontology, Ontologi DWIPA, Tourism, *Amenity*, *Ontology Enrichment*, TripAdvisor, TF-IDF, *K-Means*, K-NN, SOM, *Kappa Value*, Data Driven, *OntoQA*.

**IMPLEMENTATION OF ONTOLOGY *ENRICHMENT* METHOD WITH
AMENITY CONCEPT ON ONTOLOGY IN TOURISM DOMAIN
(case study: DWIPA Ontology)**

Shania Isyahrani

ABSTRACT

Responding to the development of Tourism in Indonesia, Ontology DWIPA III was built to represent the tourism domain in Indonesia, especially in the Regency of Bali. DWIPA III Ontology consists of *Class/subClass* Accommodation, Attraction, Regency and Event. Tourism consist of 3 main elements, namely Attraction, Accommodation and *Amenity*. Related to this, Ontology DWIPA III still has shortcomings in the *Amenity* concept that represent the tourism domain. This research uses the Ontology *Enrichment* methodology to add new concepts to the existing DWIPA ontology. The sources of the data is scraped from TripAdvisor web. *Amenity* data that has been collected is given a feature to differentiate the amenities, and the features is determined by its weighted using the Terms Frequency – Inverse Document Frequency method. The process continue by *Classifying* and comparing *K-Means*, *K-Nearest Neighbor* and *Self-Organizing Maps* algorithms. The *Kappa Value* shows that the *K-Means* algorithm has an more better accuracy declared as 0.69 than *K-Nearest Neighbor* and *Self-Organizing Maps* algorithms. Furthermore, Ontology is evaluated using Data Driven and OntoQA. The final results of this study indicate that DWIPA Ontology was successfully enriched by having 4 main *Classes*, 29 *subClasses* and 319 instances, with Richness value declared for Relationship Richness 0.33, Inheritance Richness as 32 and Attribute Richness as 79.

Kata Kunci: *Ontology, Ontology DWIPA, Tourism, Amenity, Ontology Enrichment, TripAdvisor, TF-IDF, K-Means, K-NN, SOM, Kappa Value, Data Driven, OntoQA*

DAFTAR ISI

HALAMAN PERNYATAAN ORISINILITAS	ii
HALAMAN PENGESAHAN.....	iii
UNGKAPAN TERIMAKASIH	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI.....	vi
ABSTRAK	vii
<i>ABSTRACT</i>	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
DAFTAR PERSAMAAN	xv
DAFTAR SINGKATAN.....	xvi
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian	6
1.4 Batasan Masalah.....	6
1.5 Sistematika Penelitian	7
1.6 Manfaat Penelitian.....	8
BAB II	9
TINJAUAN PUSTAKA.....	9
2.1. <i>Ontology</i>	9
2.2. <i>OWL</i>	10
2.3. <i>Ontology Editor</i>.....	11
2.3.1 <i>Protégé</i>	12
2.4. <i>Ontology Learning</i>	12
2.4.1 <i>Ontology Enrichment</i>.....	13

2.4.2 <i>Ontology Population</i>	14
2.5. <i>Ontology Evaluation</i>	16
2.5.1 <i>Data Driven</i>	16
2.5.2 <i>OntoQA</i>	16
2.6 <i>TF-IDF</i>	18
2.7 <i>Machine Learning Technique</i>	20
2.7.1 <i>K-Means</i>	20
2.7.2 <i>K-Nearest Neighbor</i>	21
2.7.3 <i>Self-Organizing Map</i>	23
2.8 <i>Kappa Value</i>	25
2.9 <i>Feature Selection</i>	25
2.10 <i>Data Analysis Tools</i>	26
2.10.1 <i>RapidMiner</i>	26
2.11 <i>Data Scraping</i>	27
2.11.1 <i>Social Media Scraping</i>	27
2.12 <i>Web Scraping Tools</i>	28
2.12.1 <i>Google-Scraper</i>	28
2.13 <i>Proses Manual, Otomatis and Semi-Otomatis</i>	28
2.14 <i>Tourism</i>	29
2.14.1 <i>Komparasi Web Tourism</i>	30
2.14.2 <i>Tourism Ontology</i>	31
2.15 <i>Penelitian Terkait</i>	32
2.16 <i>Summary</i>	35
BAB III	36
METODE PENELITIAN	36
3.1 <i>Kerangka Penelitian</i>	36
3.2 <i>Ontologi Dwipa III</i>	38
3.3 <i>Pengumpulan Data</i>	39
3.4 <i>Penentuan Fitur</i>	40
3.5 <i>Proses Pembobotan TF-IDF</i>	41

3.6 Proses Klastering	42
3.7 Proses Evaluasi.....	44
3.7.1 Data Driven.....	44
3.7.2 <i>Schema Metric OntoQA</i>	45
3.8 <i>Summary</i>	46
BAB IV	47
4.1 Pengumpulan Data.....	47
4.2 Klasifikasi <i>Terms</i>	49
4.2.1 Pemberian Fitur	49
4.2.2 Pembobotan Fitur (TF-IDF)	51
4.3 <i>Data Clustering</i>	52
4.3.1 <i>K-Means</i>	52
4.3.2 <i>K-Nearest Neighbor</i>	55
4.3.3 <i>Self-Organizing Maps</i>	58
4.4 <i>Kappa Value</i>	60
4.5 <i>Dataset Candidate</i>	61
4.6 <i>Class/Concept Detection</i>	61
4.6.1 <i>Ontology Enrichment</i>	62
4.6.2 <i>Ontology Population</i>	66
4.7 Evaluasi Ontologi.....	67
4.7.1 <i>Data Driven Evaluation</i>	67
4.7.2 <i>Schema Metrics OntoQA</i>	70
4.8 Kekurangan dari hasil penelitian	73
4.9 <i>Discussion</i>	73
BAB V	76
KESIMPULAN DAN SARAN	76
5.1 Kesimpulan.....	76
5.2 Saran	77
DAFTAR PUSTAKA	78
LAMPIRAN	83

Lampiran I: 10 Tempat Pariwisata Prioritas.....	84
Lampiran II: hasil <i>Scraping Data</i> Amenitas 5 Tempat Pariwisata Prioritas	85
Lampiran III: hasil nilai <i>Kappa Value</i> 3 Algoritma	102

DAFTAR GAMBAR

Gambar 2.1 Metode <i>Ontology Enrichment</i> [11]	13
Gambar 2.2 Metode <i>Ontology Population</i>	15
Gambar 2.3 Skema komponen OntoQA	17
Gambar 2.4 Pseudocode Algoritma <i>K-Means</i>	21
Gambar 2.5 Pseudocode Algoritma <i>K-Nearest Neighbor</i>	23
Gambar 2.6 Pseudocode Algoritma <i>Self-Organizing Maps</i>	24
Gambar 3.1 Kerangka Penelitian	37
Gambar 4.1 Tampilan <i>tools</i> Scraper	48
Gambar 4.2 <i>Flowchart</i> algoritma <i>K-Means</i>	52
Gambar 4.3 Proses pada RapidMiner menggunakan <i>K-Means</i>	53
Gambar 4.4 Cluster model yang telah dihasilkan	54
Gambar 4.5 Representasi Cluster dalam bentuk Charts.....	54
Gambar 4.6 <i>Flowchart</i> algoritma K-NN.....	55
Gambar 4.7 Proses algoritma KNN pada RapidMiner.....	56
Gambar 4. 8 Diagram proses didalam <i>Cross Validation</i>	57
Gambar 4.9 hasil pengelompokkan data menggunakan KNN	57
Gambar 4.10 <i>Flowchart</i> algoritma SOM	58
Gambar 4.11 Proses pada RapidMiner menggunakan Self-Organizing Map	59
Gambar 4.12 Proses didalam operator <i>Cross Validation</i>	59
Gambar 4.13 hasil pengelompokkan data menggunakan operator SOM.....	60
Gambar 4.14 Proses <i>Enrichment</i> DWIPA III.....	62
Gambar 4.15 Susunan <i>Class</i> baru pada ontologi DWIPA III+	63
Gambar 4.16 <i>Class Province</i>	64
Gambar 4.17 <i>Class Amenity</i>	65
Gambar 4.18 Proses Evauasi Data Driven	67
Gambar 4.19 Hasil evaluasi data driven	69

DAFTAR TABEL

Tabel 2.1 Perbandingan OWL 1 dan OWL 2.....	11
Tabel 2.2 Perbandingan Web <i>Tourism</i>	31
Tabel 3.1 Summary DWIPA III.....	39
Tabel 3.2 Contoh Data Latih 5 tempat pariwisata prioritas	41
Tabel 4.1 Hasil pengambilan data yang telah dipindahkan dalam bentuk csv.....	49
Tabel 4.2 Fitur pada setiap konsep.....	50
Tabel 4.3 Nilai <i>Kappa Value</i>	61
Tabel 4.4 Object Properties.....	66
Tabel 4.5 Hasil Relationship Richness (RR) DWIPA III+	71
Tabel 4.6 Hasil Inheritance Richness (IR) DWIPA III+	71
Tabel 4.7 Hasil Attribute Richness (AR) DWIPAIII+	72
Tabel 4.8 hasil Schema Metric OntoQA – perbandingan DWIPA III dan DWIPA III+	74

DAFTAR PERSAMAAN

Persamaan (2.1)	<i>Relationship Richness (RR)</i>
Persamaan (2.2)	<i>Inheritance Richness (IR)</i>
Persamaan (2.3)	<i>Attribute Richness (AR)</i>
Persamaan (2.4)	<i>Ontology Enrichment Score</i>
Persamaan (2.5)	<i>Term Frequency (TF)</i>
Persamaan (2.6)	<i>Inverse Document Frequency (IDF)</i>
Persamaan (2.7)	<i>Term Frequency- Inverse Document Frequency (TF-IDF)</i>
Persamaan (2.8)	<i>K-Means menggunakan Cosine Similarity</i>
Persamaan (2.9)	<i>KNN menggunakan 2-dimensional Euclidean Distance</i>
Persamaan (2.10)	<i>SOM menggunakan 1-dimensional Euclidean Distance</i>
Persamaan (2.11)	<i>Menghitung nilai Kappa Value</i>
Persamaan (2.12)	<i>Persamaan Kappa Value ditulis dengan simbol</i>

DAFTAR SINGKATAN

3A	<i>Atraksi, Aksesibilitas dan Amenitas</i>
OWL	<i>Web Ontology Language</i>
OntoQA	<i>Ontology Quality Analysis/Assesment</i>
RR	<i>Relationship Richness</i>
IR	<i>Inheritance Richness</i>
AR	<i>Attribute Richness</i>
TF	<i>Term Frequency</i>
IDF	<i>Inverse Document Frequency</i>
K-NN	<i>K-Nearest Neighbor</i>
SOM	<i>Self-Organizing Map</i>
YALE	<i>Yet Another Learning Environment</i>