Unveiling Online Perception toward "1 Malaysia" over Time: A Social Impact Analysis

By **Sunway University**

Dr Lim Tong Ming
Dr Wong Siew Fan
Dr Ngeow Yeok Meng
Mr Vicky Lim Min How
30 Oct. 2012, 2.00PM @ SKMM HQ

Objectives

▸ The objectives of this research are:

▸

▸ To understand the impact of "1Malaysia" through an analysis of citizens' online perception toward the concept

▸

▸ To develop a prototype that would provide an ontological representation of citizens' perception toward '1Malaysia'

Research Description

▸ The '1Malaysia' concept has the goals of fostering harmony and unity among people of different races and religions as well as enforcing governance and efficiency.

▸ As a top-down, nation-wide campaign, '1Malaysia' is believed to have created certain impacts among citizens of Malaysia

▸ It is both timely and important to study the impact of '1Malaysia' on its citizens. Such impact can be unveiled through an analysis of citizens' perception toward the concept over time.

▸ The analysis helps evaluating the effectiveness of '1Malaysia' and determining if any intervening plan is needed to further enforce the concept among the citizens.

▸ A popular source to express feeling, opinions, and thoughts, including their perception toward '1Malaysia' is web blogs.

▸ This project therefore intends to establish an understanding of the impact '1Malaysia' has on citizens through an analysis of their perception published on web blogs.
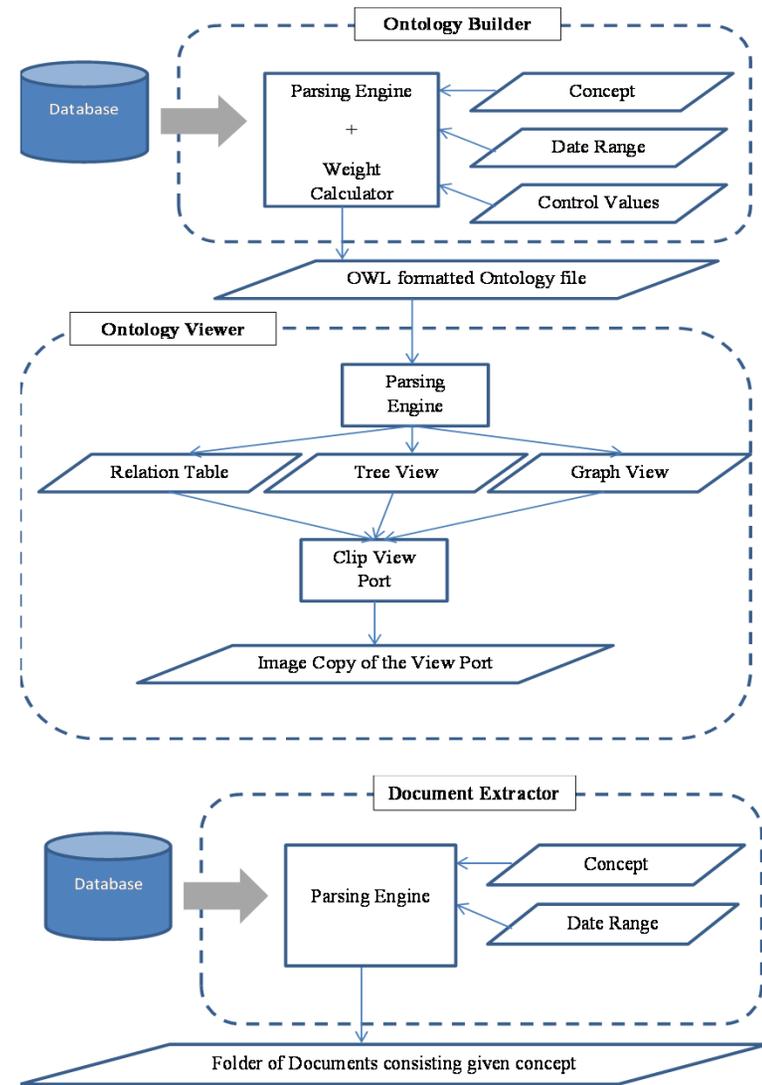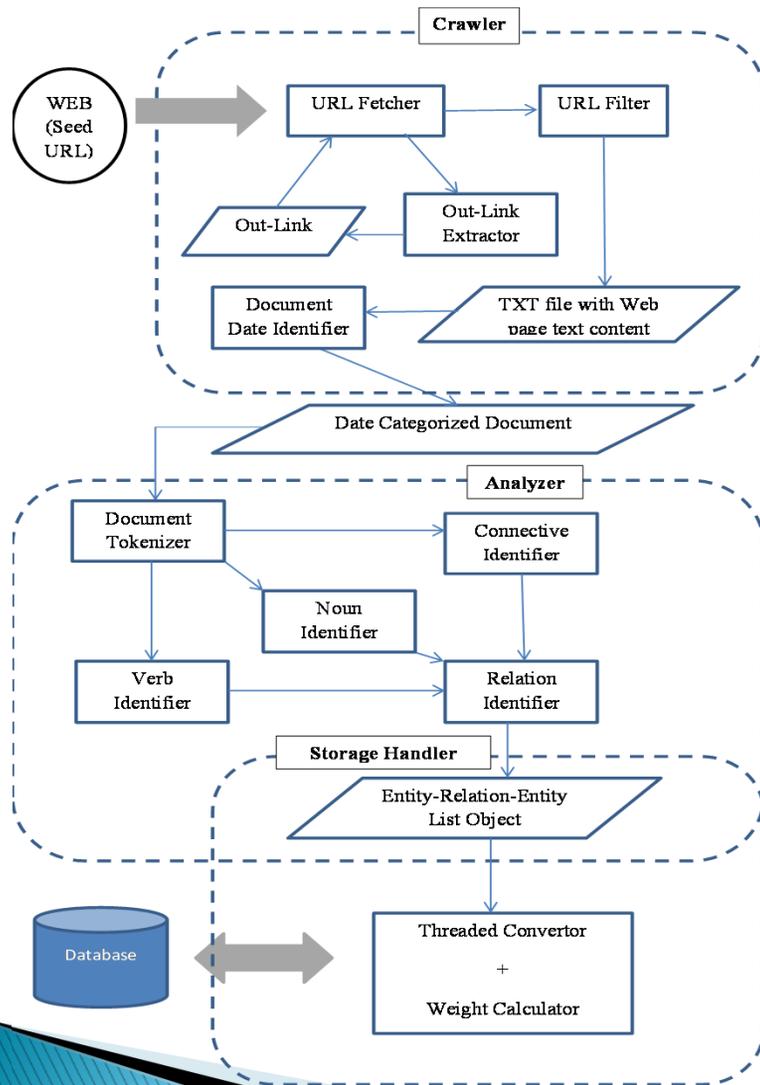
Outlines of Presentation

▸ The Prototype: Temporal Driven Text-2-Onto
▸ The Social Impact Analysis on "1Malaysia"
▸ Q&A

The Prototype: Temporal Driven Text-2-Onto

- The prototype is capable of constructing a temporal based ontological (tree) representation of a specific 'concept' from a set of texts crawled
- The prototype consists of the following modules
  - Crawler
  - Analyzer
  - Storage Handler
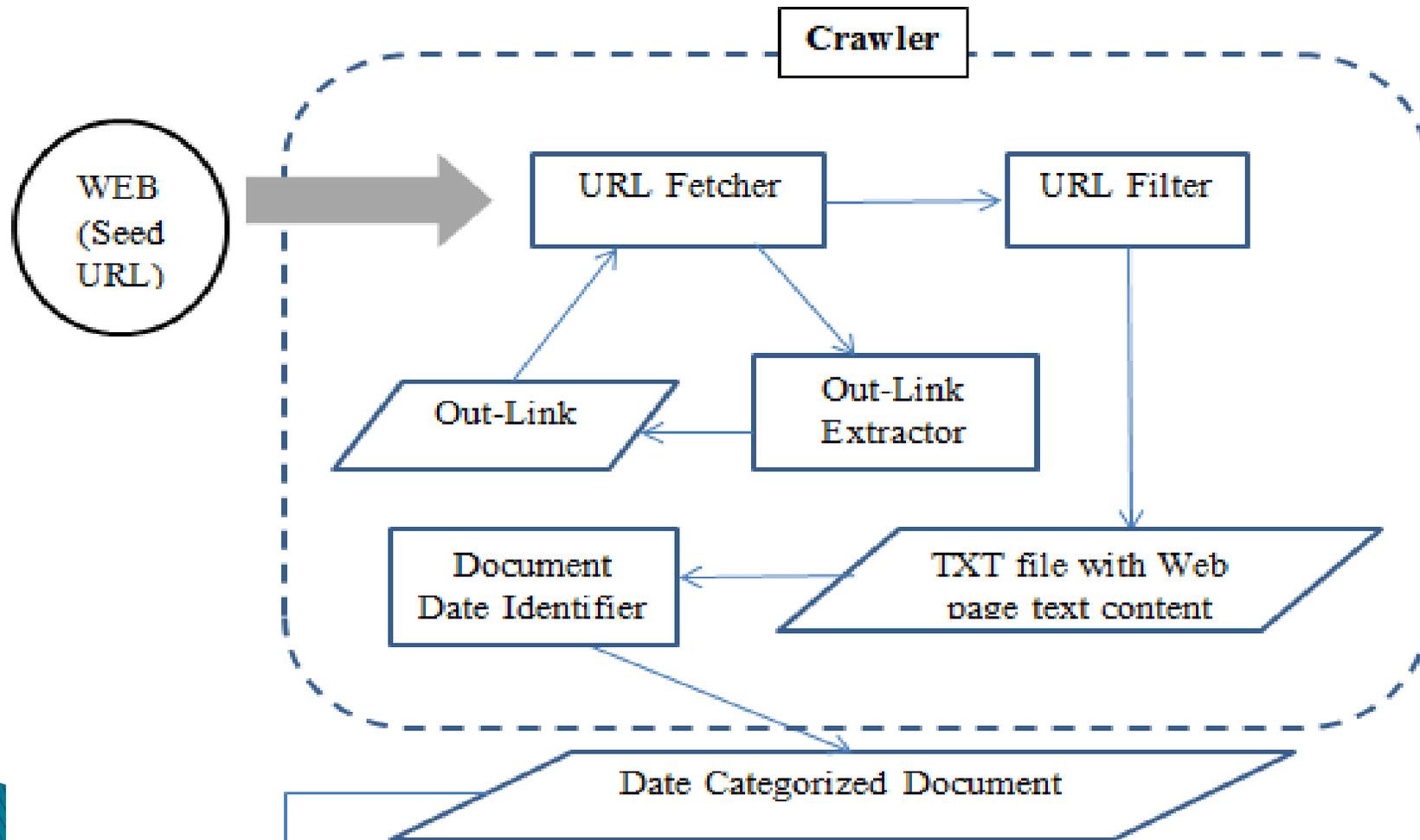 - Ontology Builder
  - Ontology Viewer
  - Document Extractor

The Architecture

Crawler

WEB (Seed URL) → URL Fetcher → URL Filter

Out-Link ← Out-Link Extractor

Document Date Identifier ← TXT file with Web page text content

Date Categorized Document

Analyzer

Document Tokenizer → Connective Identifier

Noun Identifier

Verb Identifier → Relation Identifier

Storage Handler

Entity-Relation-Entity List Object

Database ↔ Threaded Convertor + Weight Calculator

Ontology Builder

Database → Parsing Engine + Weight Calculator ← Concept

Date Range

Control Values

OWL formatted Ontology file

Ontology Viewer

Parsing Engine

Relation Table | Tree View | Graph View

Clip View Port

Image Copy of the View Port

Document Extractor

Database → Parsing Engine ← Concept

Date Range

Folder of Documents consisting given concept

Crawler

- Functionality
  - Craw web document
  - Filter by URL
  - Simple Date Detection
- Usage
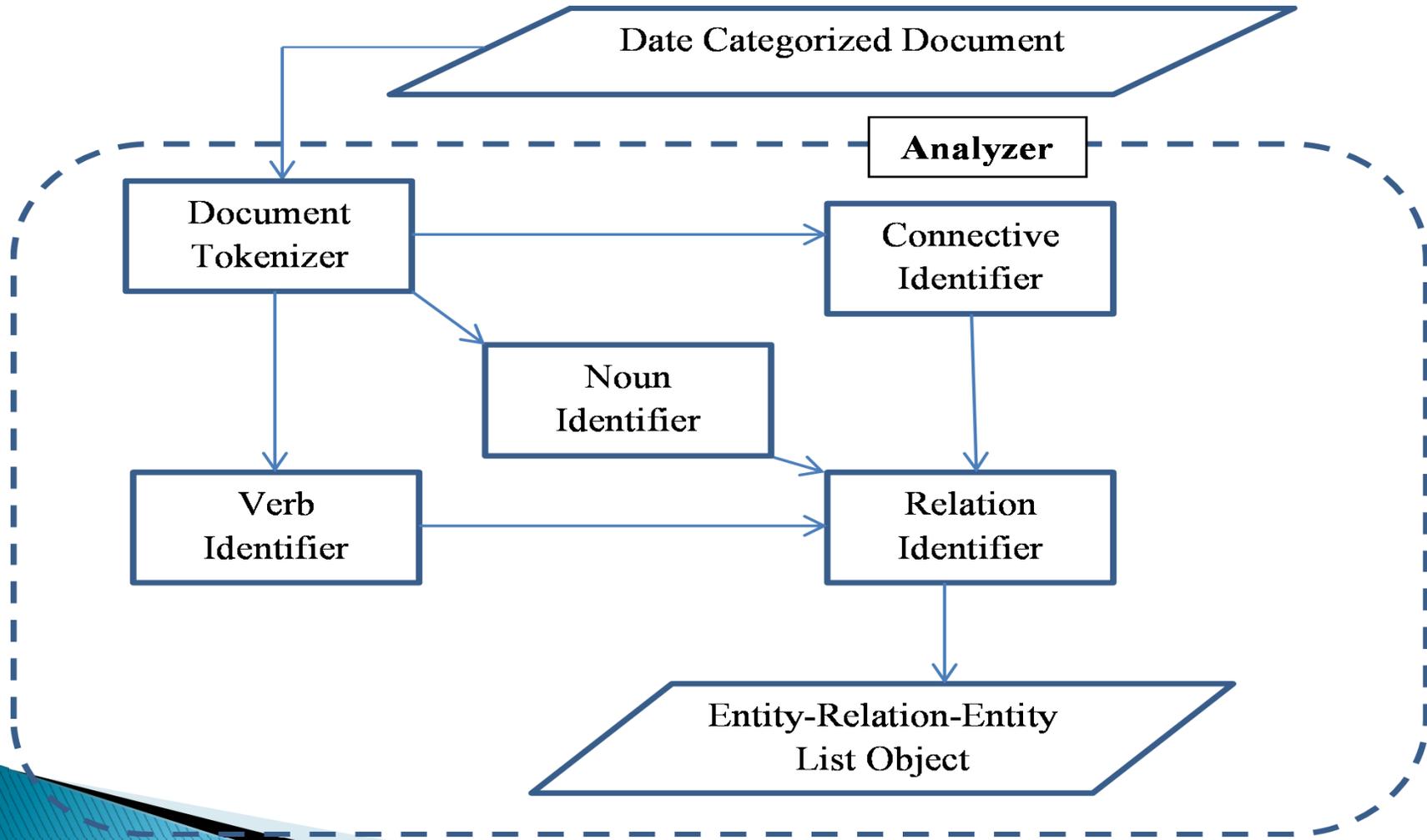 - Collect/Gather information sources

Crawler

Analyzer

▸ Functionality
- Disassemble well written text into grammatical parts
- Identify noun and verb phrases
- Re-assemble the phrases into Entity-Relation-Entity (ERE) component together with document references as tetrad
  - E.g. "Mommy"(Noun)-"love"(Verb)-"Hiking"(Noun)-"Document Reference"

▸ Usage
- Dissect document into data component

Analyzer

Date Categorized Document

Analyzer

Document Tokenizer

Connective Identifier

Noun Identifier

Verb Identifier

Relation Identifier

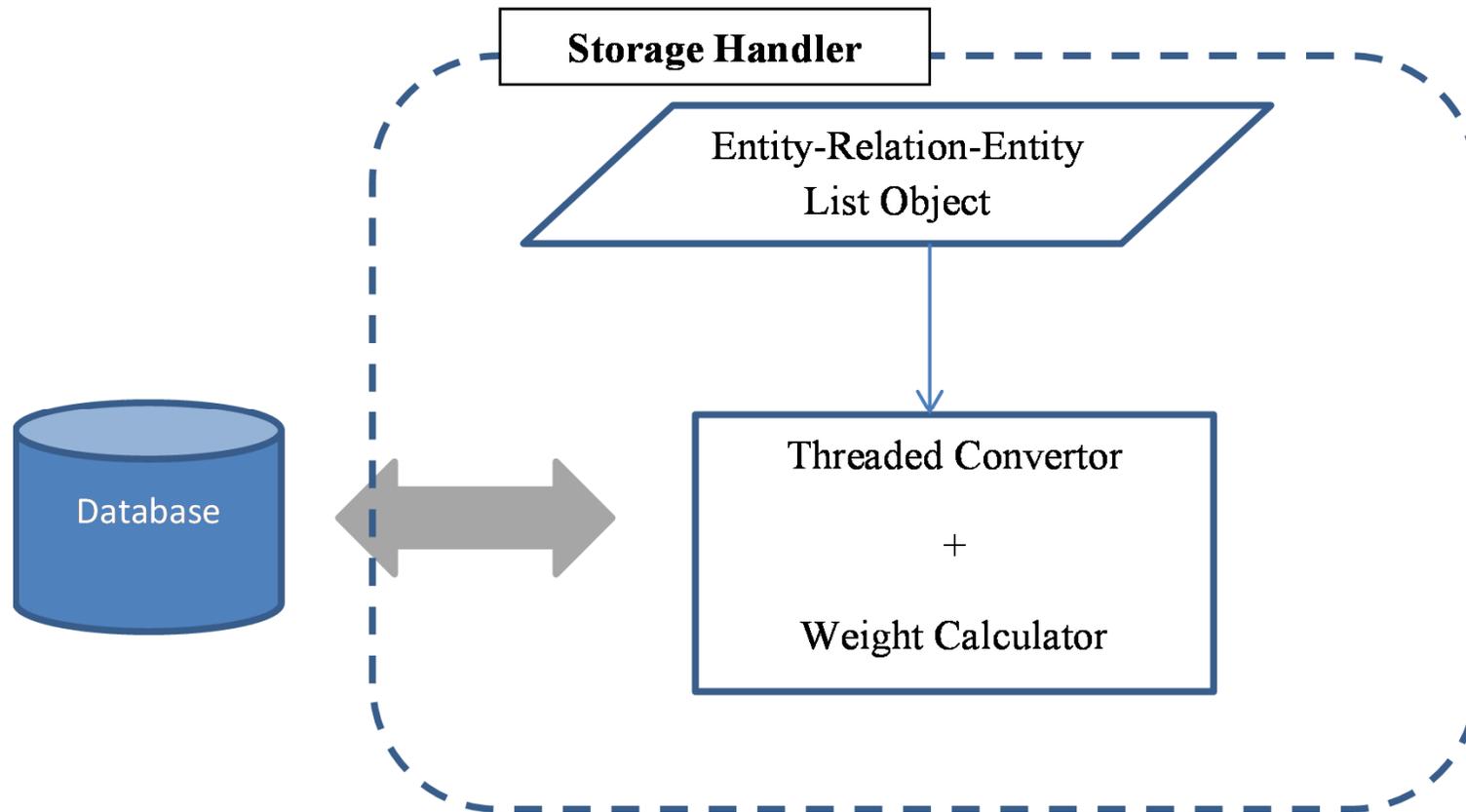Entity-Relation-Entity List Object

Storage Handler

- Functionality
  - Parse the tetrad from previous sub-module
  - Convert/Translate into database
  - Calculate and Update the base score
- Usage
  - To calculate and store the basic information for ontology building

Storage Handler

Storage Handler

Entity-Relation-Entity List Object

Database

Threaded Convertor
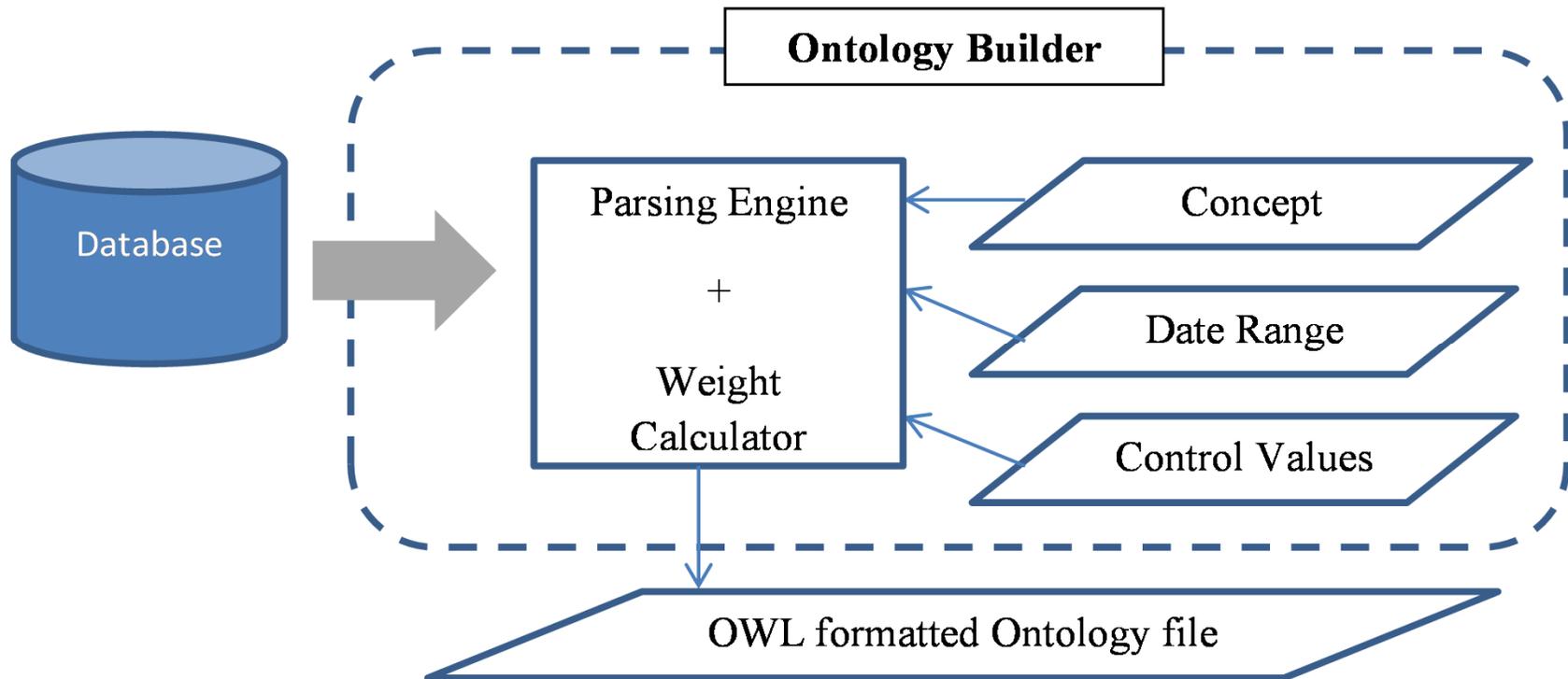
+

Weight Calculator

Ontology Builder

- Functionality
  - Get user input:
    - Core/Start concept
    - Date Range
    - Control Value (eg. Level depth, sibling number)
  - Parse Database with given input/parameter
  - Output OWL formatted file for constructed Ontology tree
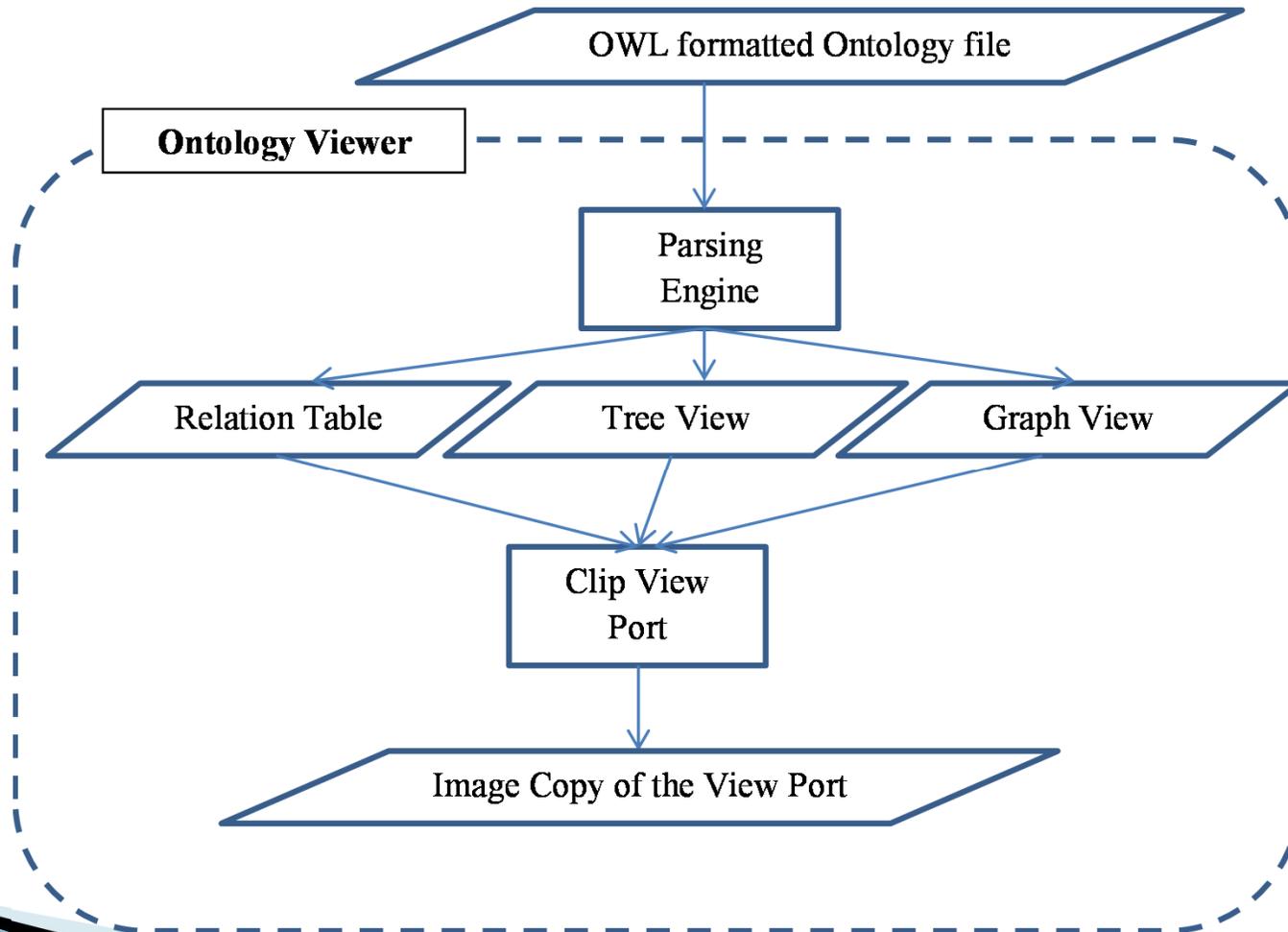- Usage
  - Build Ontology tree

Ontology Builder

Ontology Viewer

▶ Functionality
- ◦ Parse OWL formatted file (General/Customized)
- ◦ Construct
  - • Tree Visualization
  - • Relation Table
- ◦ Snap part of Visualized Ontology tree

▶ Usage
- ◦ Visualize generated OWL formatted file

Ontology Viewer

OWL formatted Ontology file

Ontology Viewer

Parsing Engine

Relation Table   Tree View   Graph View

Clip View Port

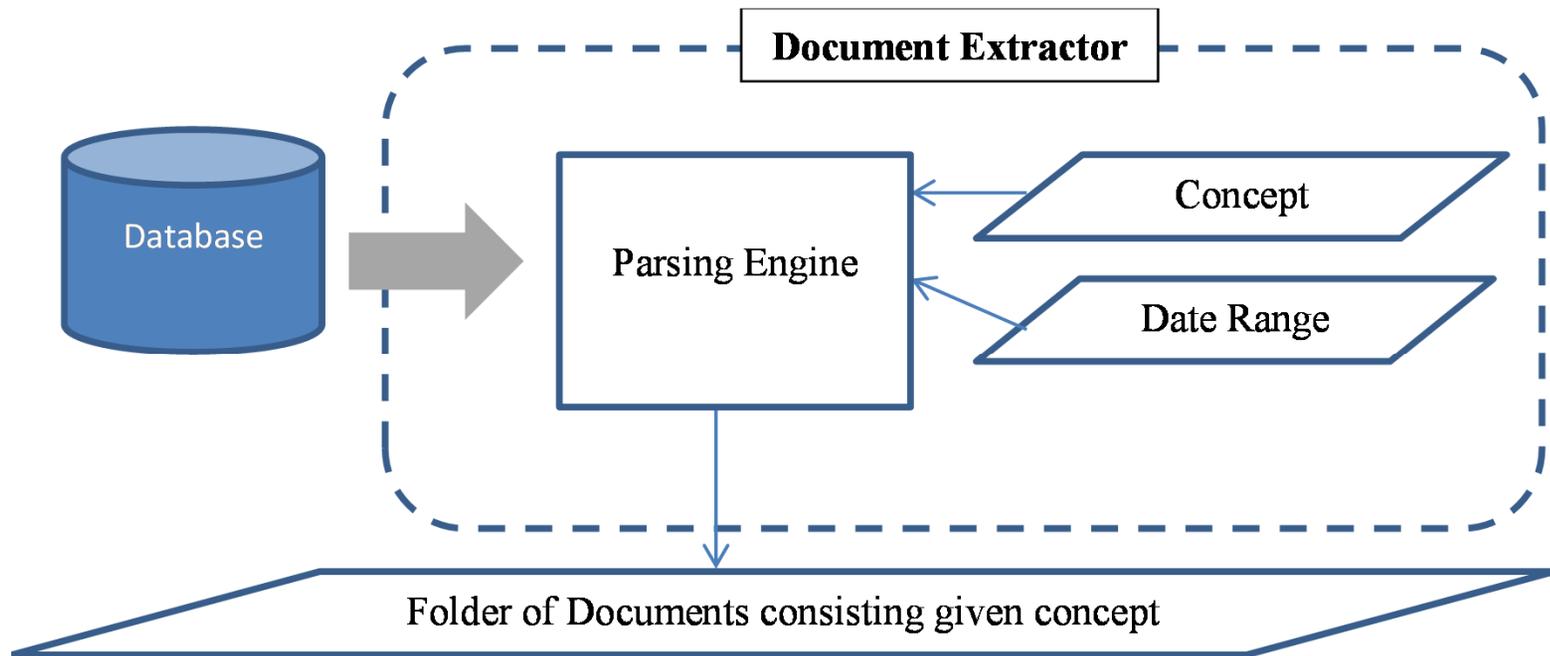Image Copy of the View Port

Document Extractor

- Functionality
  - Gather input:
    - Core/Start concept
    - Date Range
    - Output folder
  - Parse Database with given input/parameter
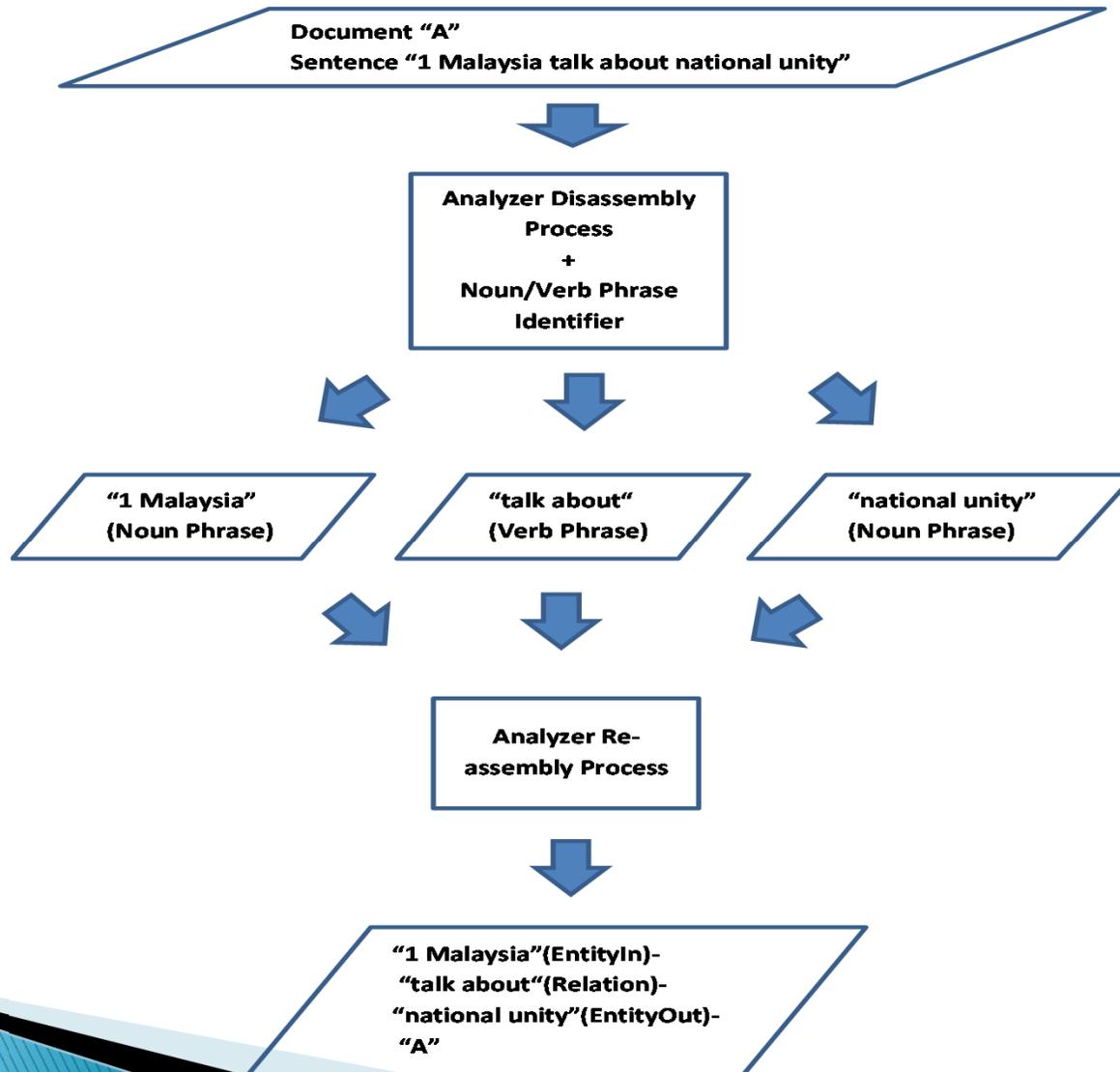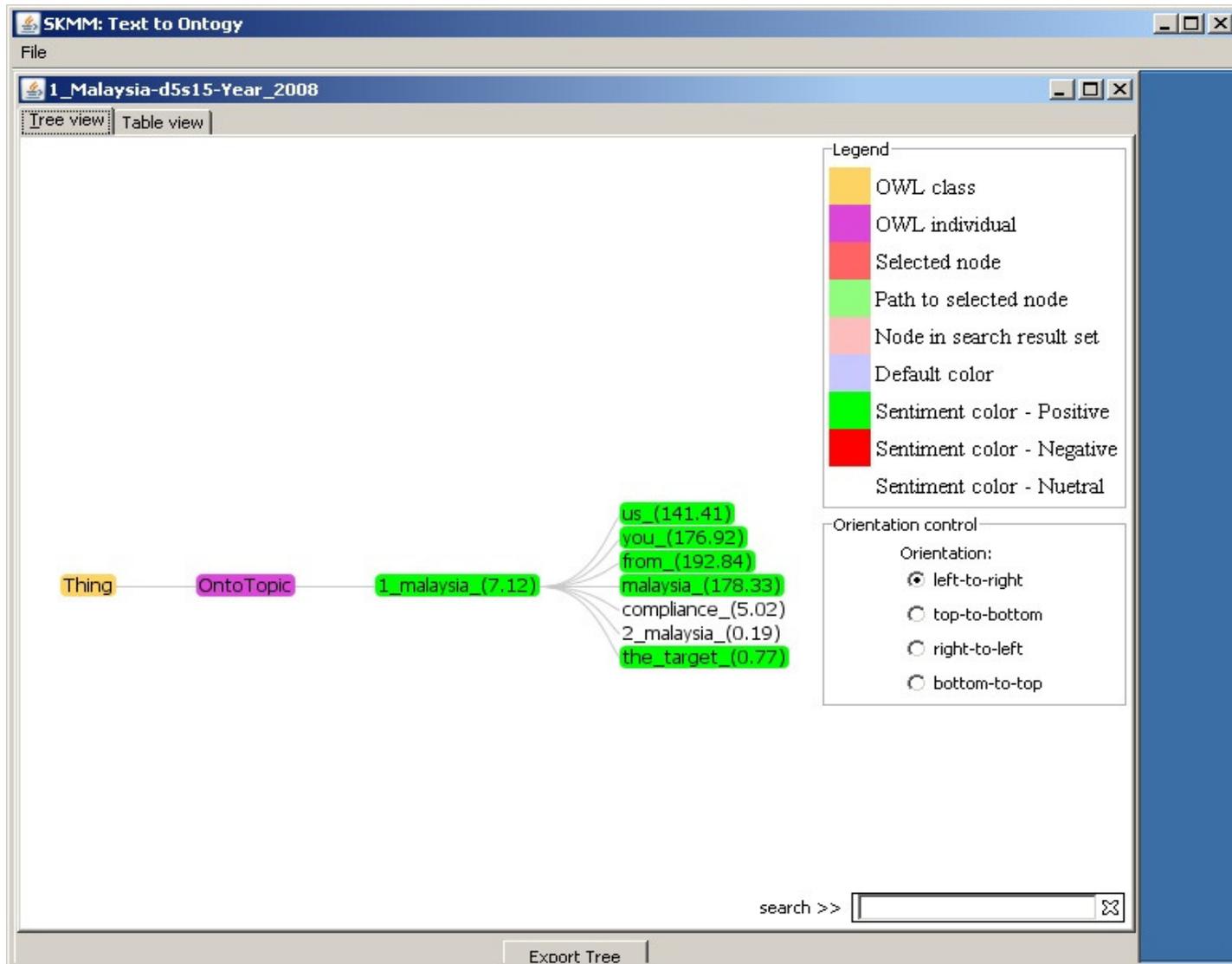 - Extract related documents into selected folders

- Usage
  - Extract documents which contribute to interested Concept Entity in the Visualized Ontology Tree
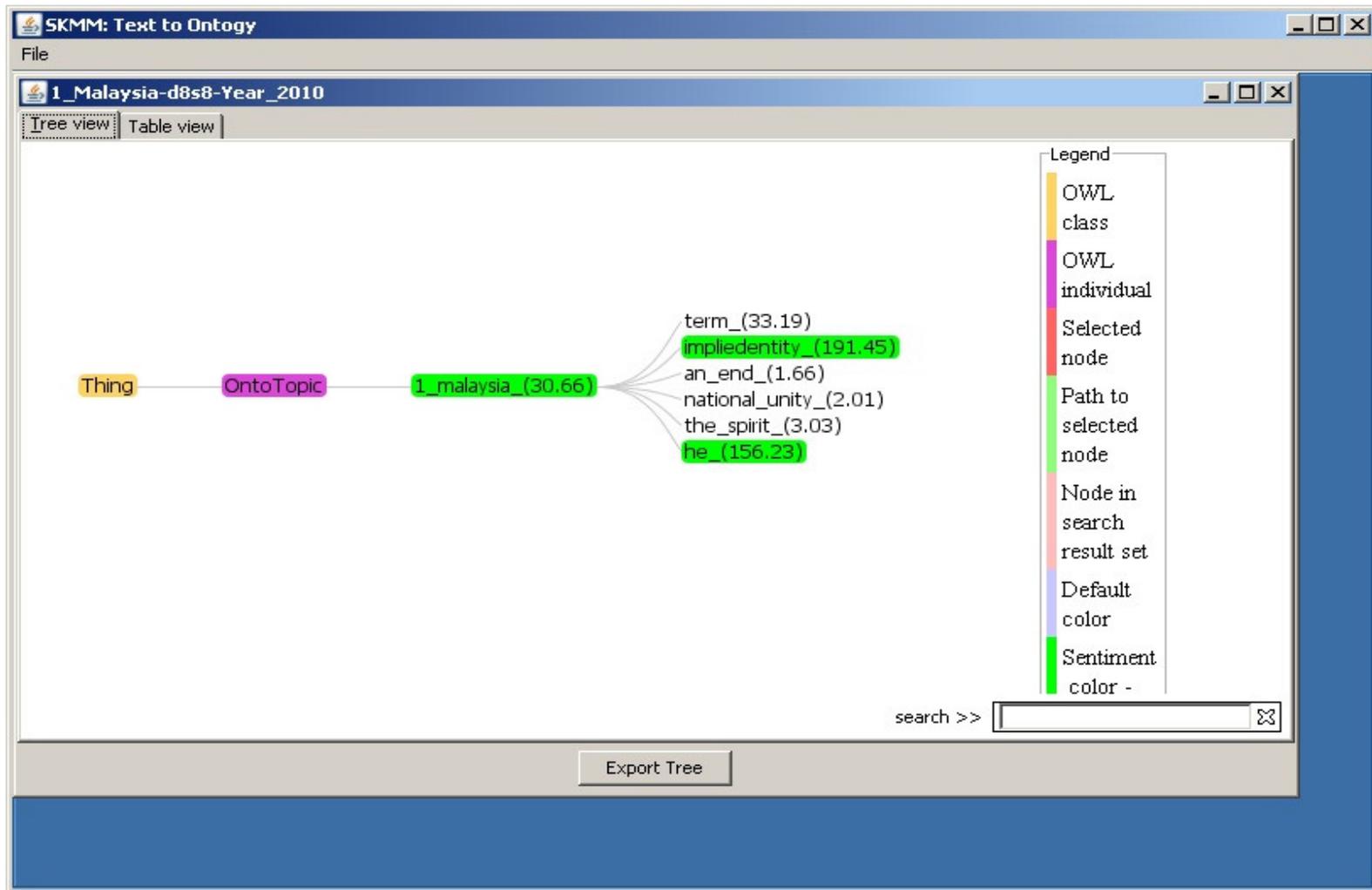
Document Extractor

Document Extractor

Database → Parsing Engine ← Concept

Parsing Engine ← Date Range

Folder of Documents consisting given concept

How it Work?

Document "A"
Sentence "1 Malaysia talk about national unity"

↓

Analyzer Disassembly Process
+
Noun/Verb Phrase Identifier

↙ ↓ ↘

"1 Malaysia"
(Noun Phrase)

"talk about"
(Verb Phrase)

"national unity"
(Noun Phrase)

↘ ↓ ↙

Analyzer Re-assembly Process

↓

"1 Malaysia"(EntityIn)-
"talk about"(Relation)-
"national unity"(EntityOut)-
"A"

Ontology Model for Concept "1 Malaysia" on October to December 2008.

Ontology Model for Concept "1 Malaysia" on July to September 2010.

The Social Impact Analysis on "1Malaysia"

- To identify the perceived social impact of the concept of One Malaysia since 2008 - 2012

i. To show ontology trees of massive data on the concept of One Malaysia.

ii. To examine the perceptions of bloggers towards the concept of One Malaysia via textual analysis.

iii. To determine the support or rejection of the concept and why.

iv. To measure the perceived effectiveness of the concept via numeric and textual analysis.

Analysis of concept weightage for four key concepts in ontology trees 2008-2012 (based on selected quarters)

Time Frame	Frequency of keyword 'People'	Frequency of keyword 'Right'	Frequency of keyword 'Nation'	Frequency of keyword 'Action'
Jul-Sep 2008	2510/2591/2601	912/1519/1610	610/715/1556	463/674
Oct-Dec 2008	4276/4353/4460	1750/2890/3145	1063/1216/3273	867/1162
Apr-Jun 2009	3341/3407/3431	1396/2561/2737	772/919/2690	833/1062
Jul-Sep 2010	5177/5314	2088/3303	1233/1518	1607/2119
Oct-Dec 2011	5484/5631	1679/3598	1258/1896	695/1422
Jan-Mar 2012	2381/3906	3670/8575	11769/12168	1659/3456

Indicator of Numeric Data in Frequency: Exacted key concept/with 's'/with any tailing

Data selection

▸ We conducted data selection based on highest frequency of four selected keywords to indicate the most frequently discussed concepts in association with the One Malaysia concept

Exclusion of elements

- Discourse by politicians, both ruling and opposition parties, thus ruling out the top-down(authoritative-defined)concept of One Malaysia interpreted by the authorities
- Media write-ups by journalists, thus ruling out the reports that feature authority-defined political discourse.
- Official or unofficial websites that may be affiliated to the ruling or opposition parties.
- Repetitive text found in different blogs, thus excluding mirror sites and technical counting produce errors.

Inclusion of elements

- Comprehensible personal blogs written in English.
- Text that are reflected of bloggers' view, perception, attitude, behavior, so that we are aware of their cognitive, affective and psychomotor domains they have experienced in writing.

Methodology

▸ **Interpretation of patterns in numeric data**
- Use the ontology tree to capture the frequency count and weightage of the concept of four identified keywords, i.e. people, nation, rights, action.
- The four concepts repeatedly appeared at high frequency, making them significant when selecting text for further interpretive analysis.
- The ontology tree showcases patterns and tendency of how bloggers discussed the ideas of One Malaysia

Methodology

- Interpretation of meaning in text
  - We make use of the above four items to see how bloggers actually wrote about the concept of One Malaysia. Here we focus on meaning and interpretation using textual approach, thus avoiding subjective interpretation of text based on themes, categories and framing analysis.
  - Item of analysis is bloggers' text (sentence, paragraph and passage depending on its content)

Analysis and Remarks on PEOPLE

- Theme No.1: PEOPLE (sample text, perception: rejection)
- Theme No.1: PEOPLE (sample text, perception: acceptance)
- Theme No.1: PEOPLE (sample text, perception: neutral)
- Theme No.1: PEOPLE (sample text, understanding, attitude)

Analysis and Remarks on RIGHTS

- Theme No.2: RIGHTS (understanding: puzzling)

Analysis and Remarks on NATION

- Theme No.3: NATION (perception: rejection)

Q&A