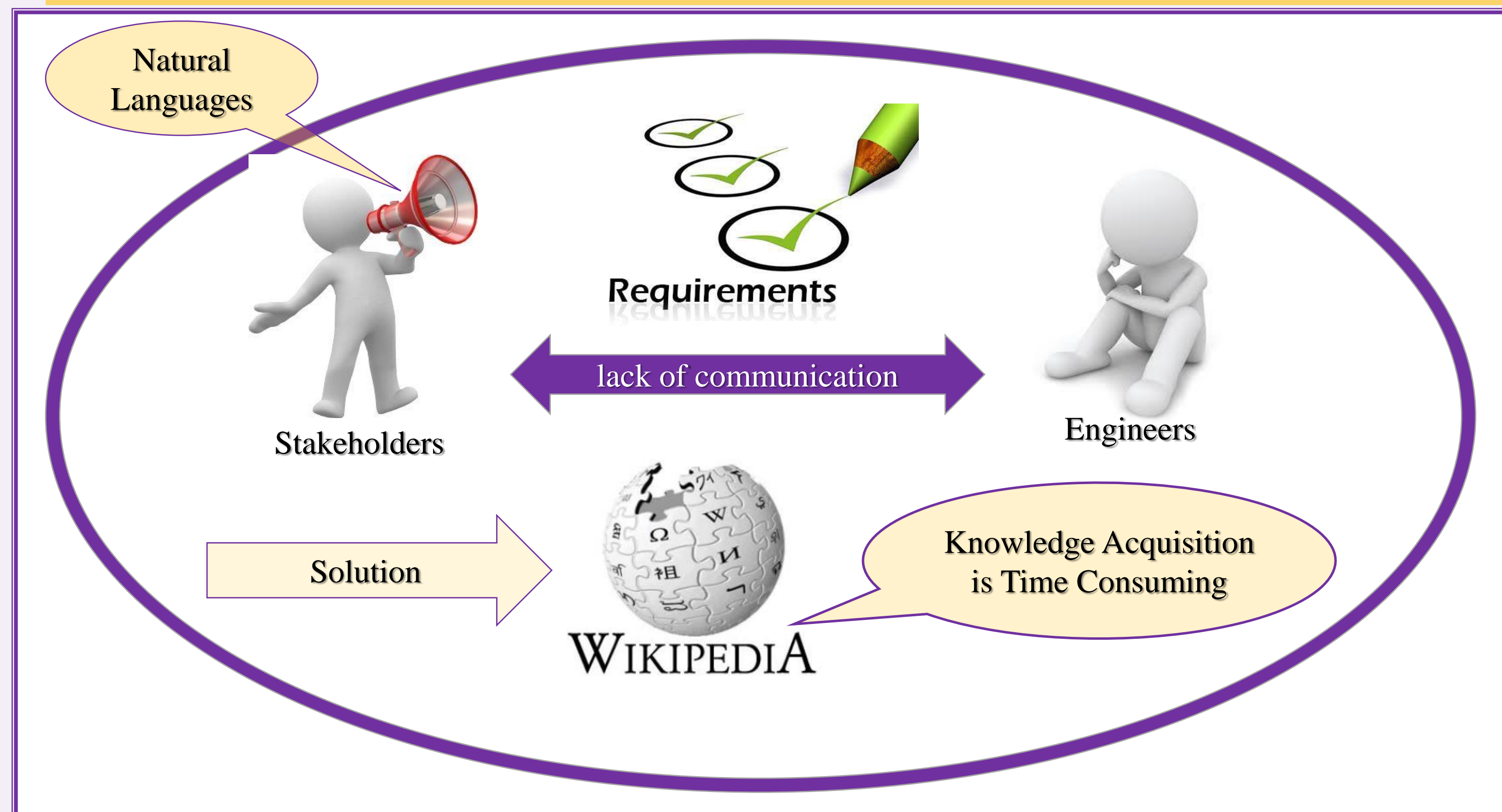# Improving the Efficiency of the Knowledge Acquisition Process in Requirements Engineering

Danissa Rodriguez, Doris Carver, Anas Mahmoud
Division of Computer Science and Engineering Louisiana State University Baton Rouge, Louisiana 70803

ABSTRACT

Natural Languages

Requirements

Stakeholders — lack of communication — Engineers

Solution

WIKIPEDIA

Knowledge Acquisition is Time Consuming

INTRODUCTION

Software engineering (SE) defines processes for the development, operation and maintenance of software.

IEEE → "The application of a systematic, disciplined, quantifiable approach to the development, operation, and maintenance of software; that is, the application of engineering to software".

The first step in the SE process for creating a software system is to determine the requirements.

From the requirement elicitation phase, where software engineers and stakeholders elicit requirements from the stakeholders.
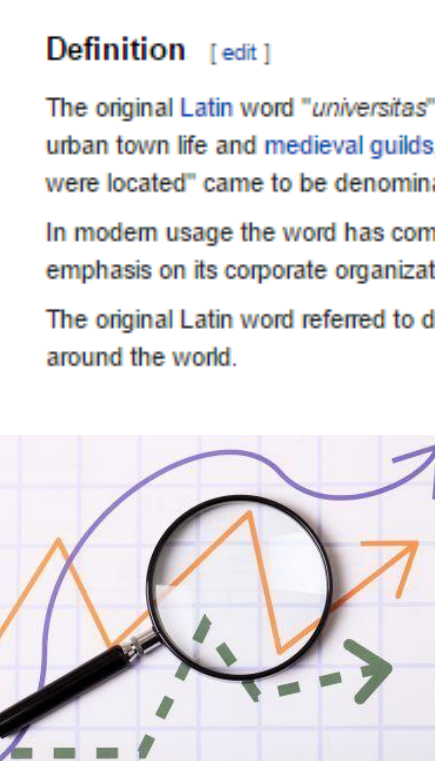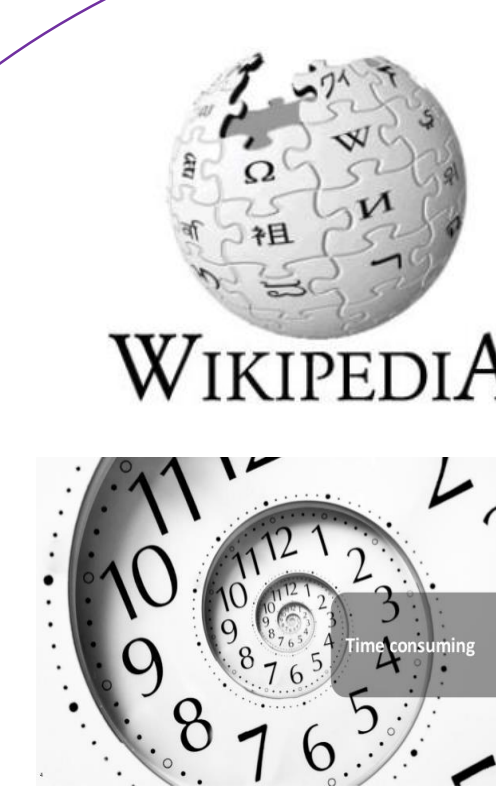
Requirements Engineering

Inception — Elicitation — Elaboration

Negotiation — Specification — Validation
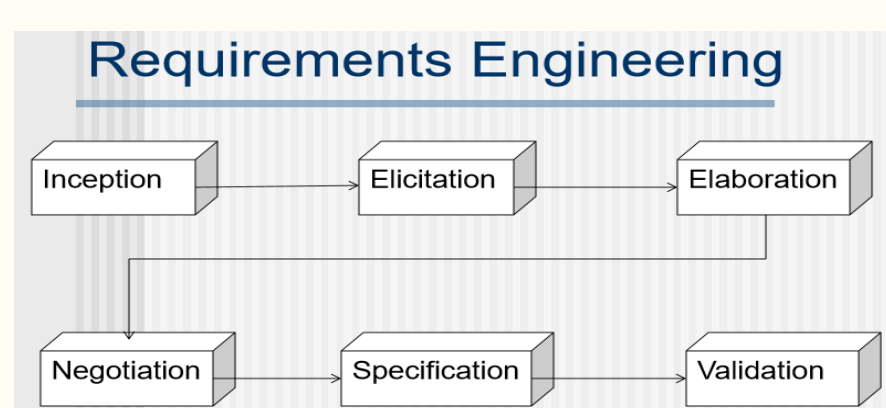
However, there is a gap in the Requirement Engineering (RE) process because of a lack of communication that occurs between stakeholders and software engineers.
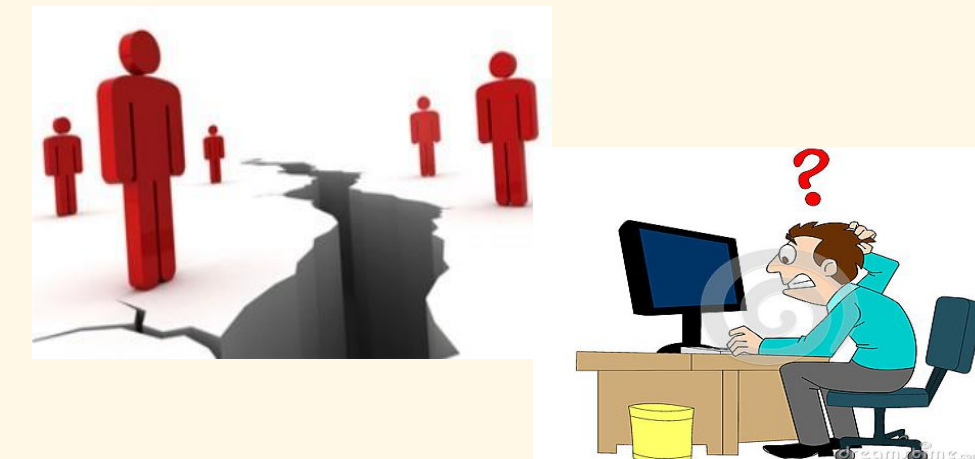
Requirements are typically expressed in natural language, and they are linked to the application area which may not to be in a developer's area of expertise.

SOLUTION

Mahmoud and Carver describe a structure to decrease the gap in generating the requirements utilizing Wikipedia-based natural language
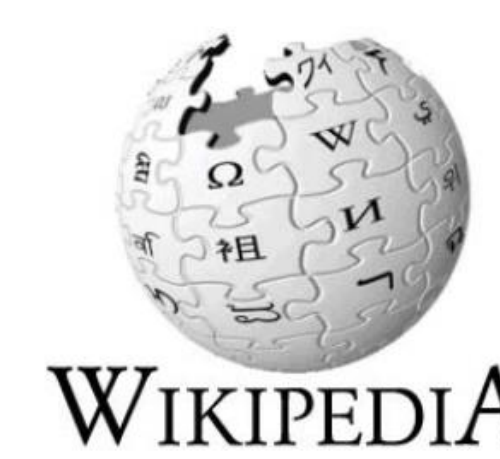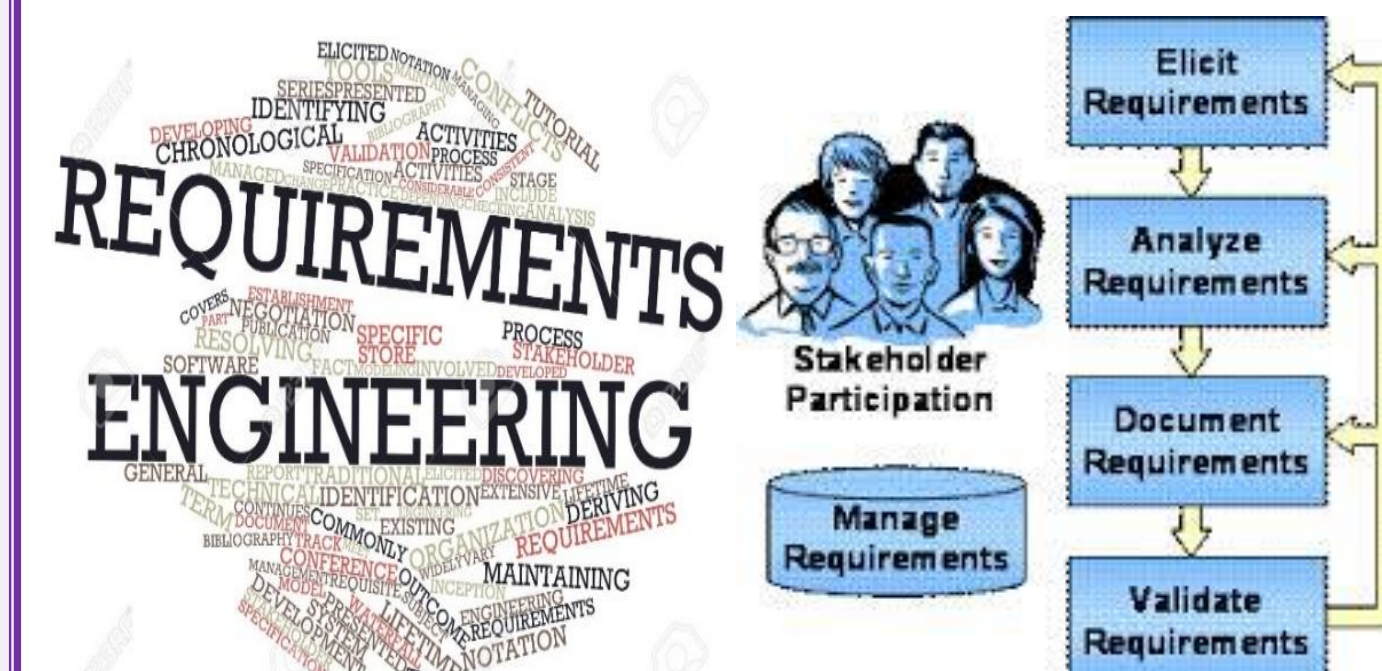
WIKIPEDIA

Knowledge acquisition for Wikipedia

BACKGROUND

Requirement Engineering

RE is the gathering and analyzing of user expectations for the software system.

REQUIREMENTS ENGINEERING

Elicit Requirements → Analyze Requirements → Document Requirements → Validate Requirements

Stakeholder Participation

Manage Requirements

Knowledge Acquisition from Wikipedia
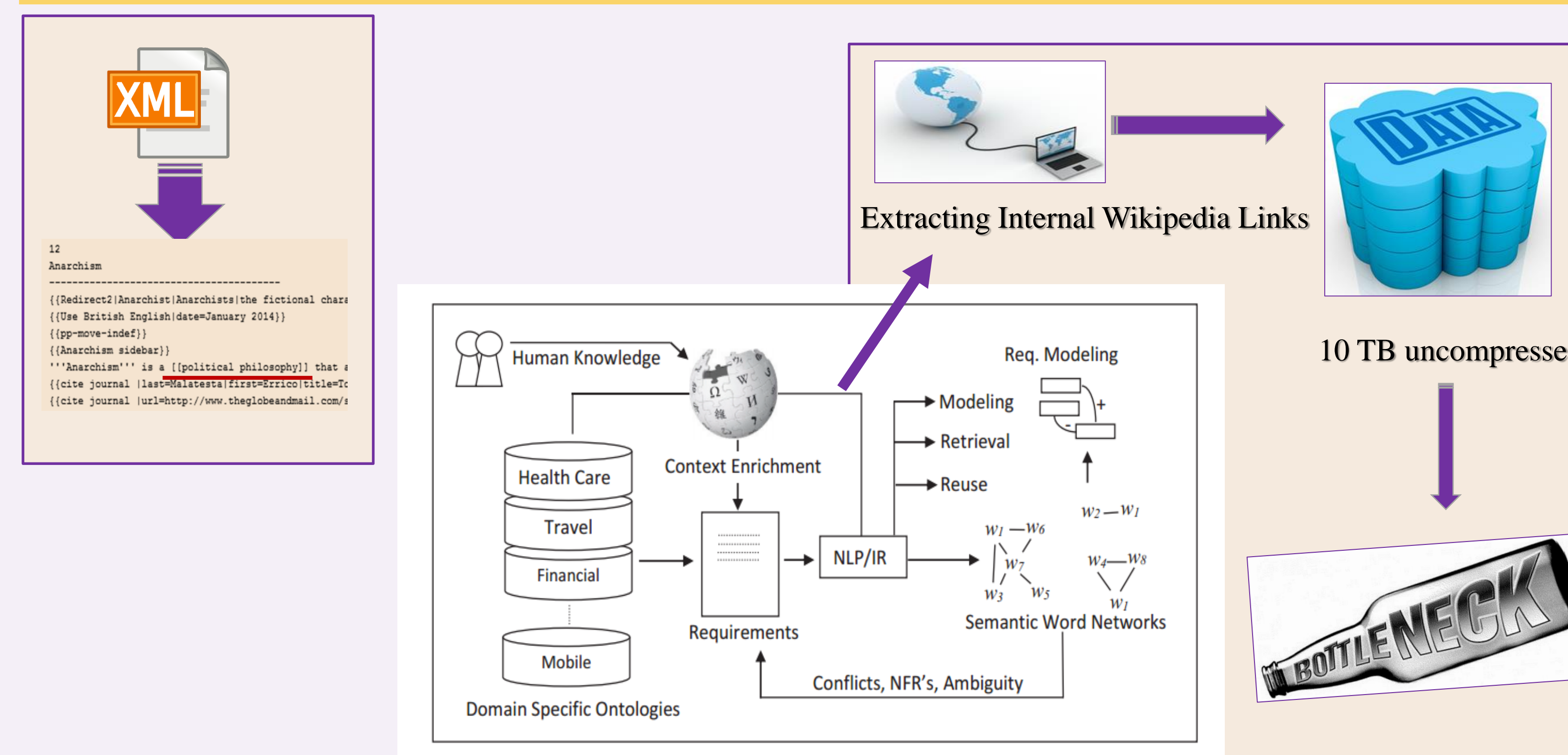
WIKIPEDIA

Extracting information from a Wikipedia corpus. It involves the process of extracting, structuring and organizing knowledge from one source

Domain Specific Ontologies

ONTOLOGY

Ontology is the specification of a conceptualization. Domain ontologies have value for facilitating the understanding between stakeholders though the establishment of a common terminology.
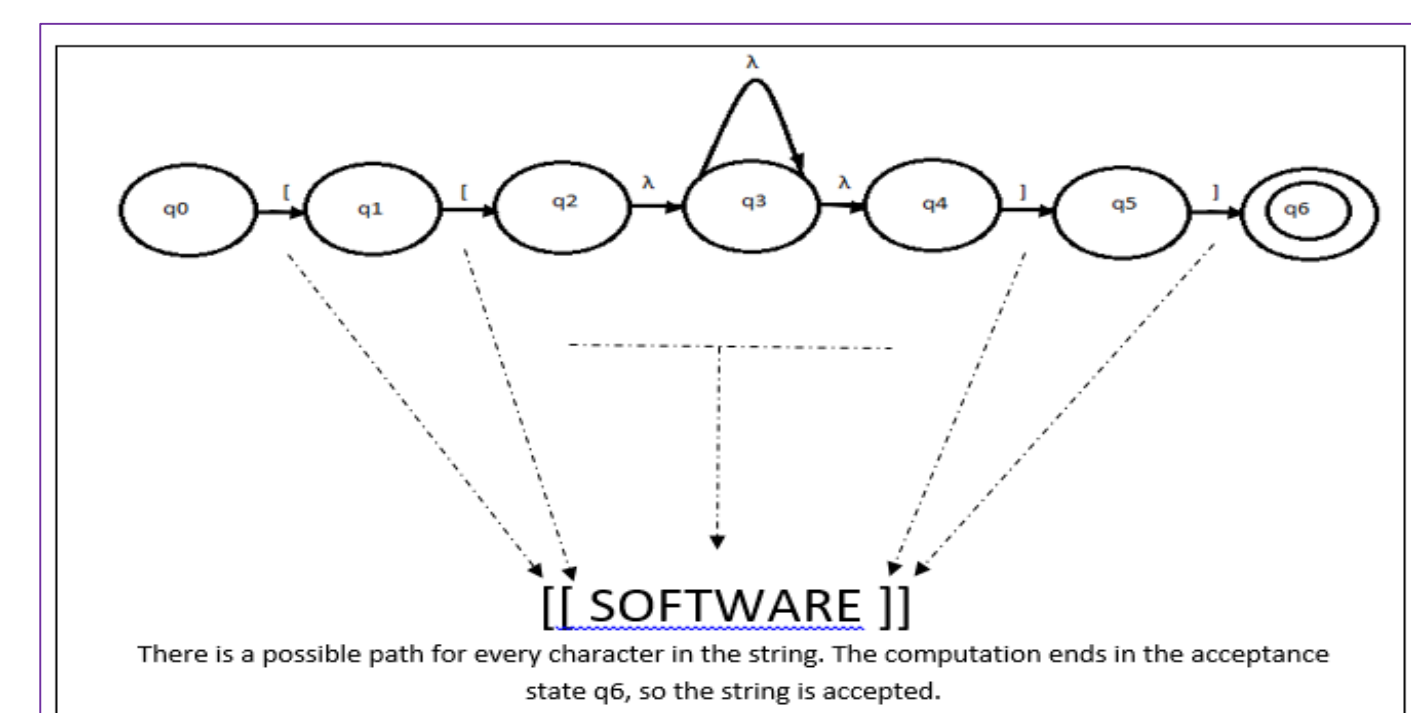
FRAMEWORK

XML

Extracting Internal Wikipedia Links

DATA

10 TB uncompressed

BOTTLENECK

Human Knowledge

Health Care / Travel / Financial / Mobile

Domain Specific Ontologies

Context Enrichment

NLP/IR

Requirements

Req. Modeling — Modeling, Retrieval, Reuse

Semantic Word Networks

$W_1 - W_5$ $W_2 - W_3$
W_3 $W_4 - W_5$
W_3 W_3

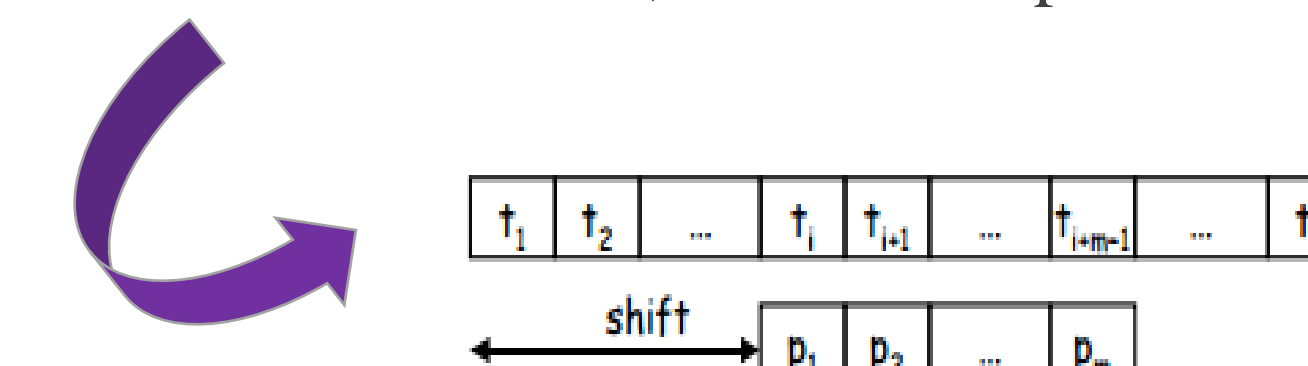Conflicts, NFR's, Ambiguity

METHODS

Regular Expression

A regular expression is a sequence of characters that define a search pattern, mainly for use in pattern matching with strings, or string matching.

[[SOFTWARE]]

There is a possible path for every character in the string. The computation ends in the acceptance state q6, so the string is accepted.

String Search Algorithms (SSAs)

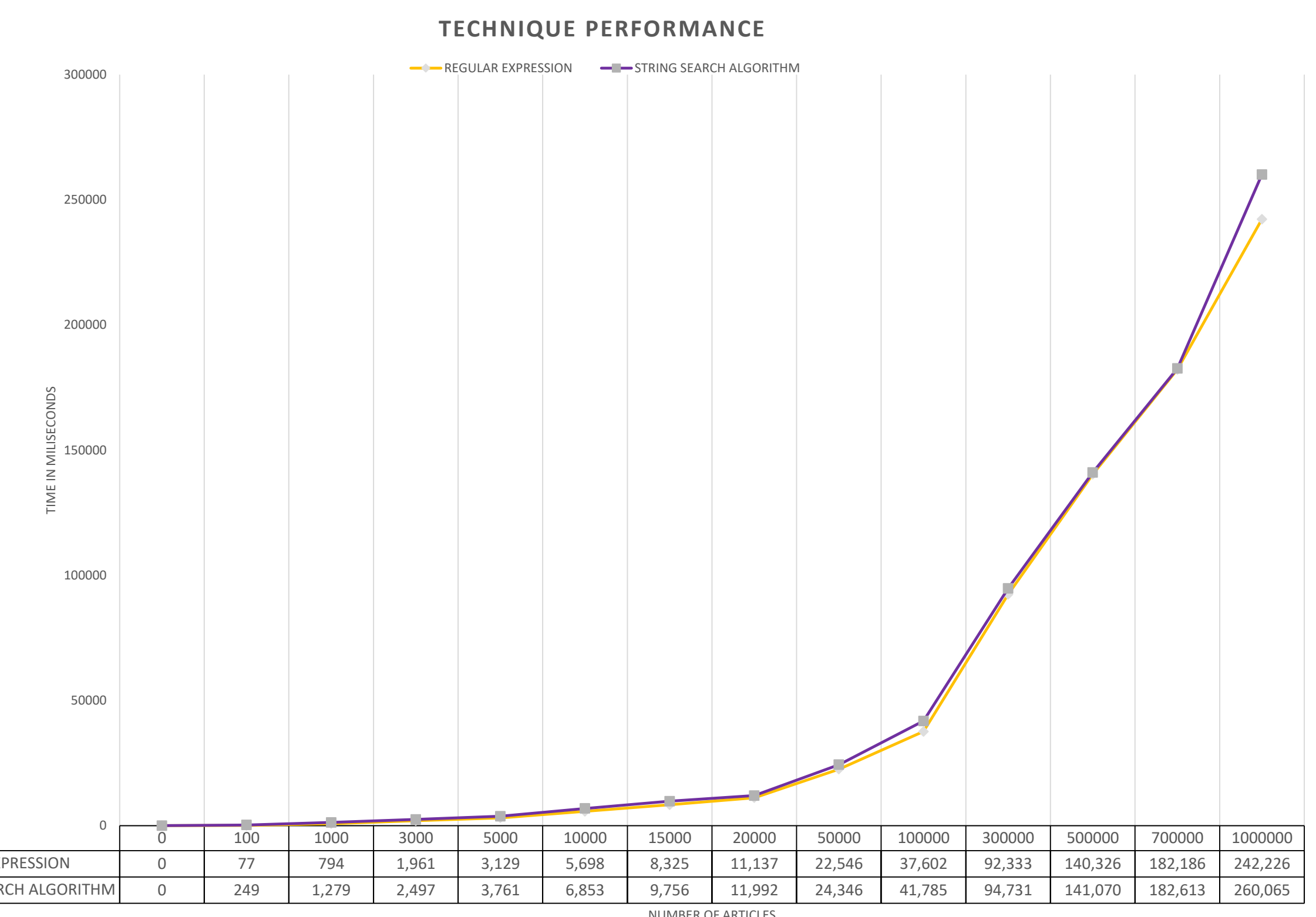SSAs are used to locate one or several strings within a large string.

- **Naive Search Algorithm**: The Naive search algorithm provides a simple solution. It is in fact the most straightforward solution. If there is a need to find a pattern P within a text T, the algorithm compares P with T at position 1, and in case of a mismatch, it shifts the position of P
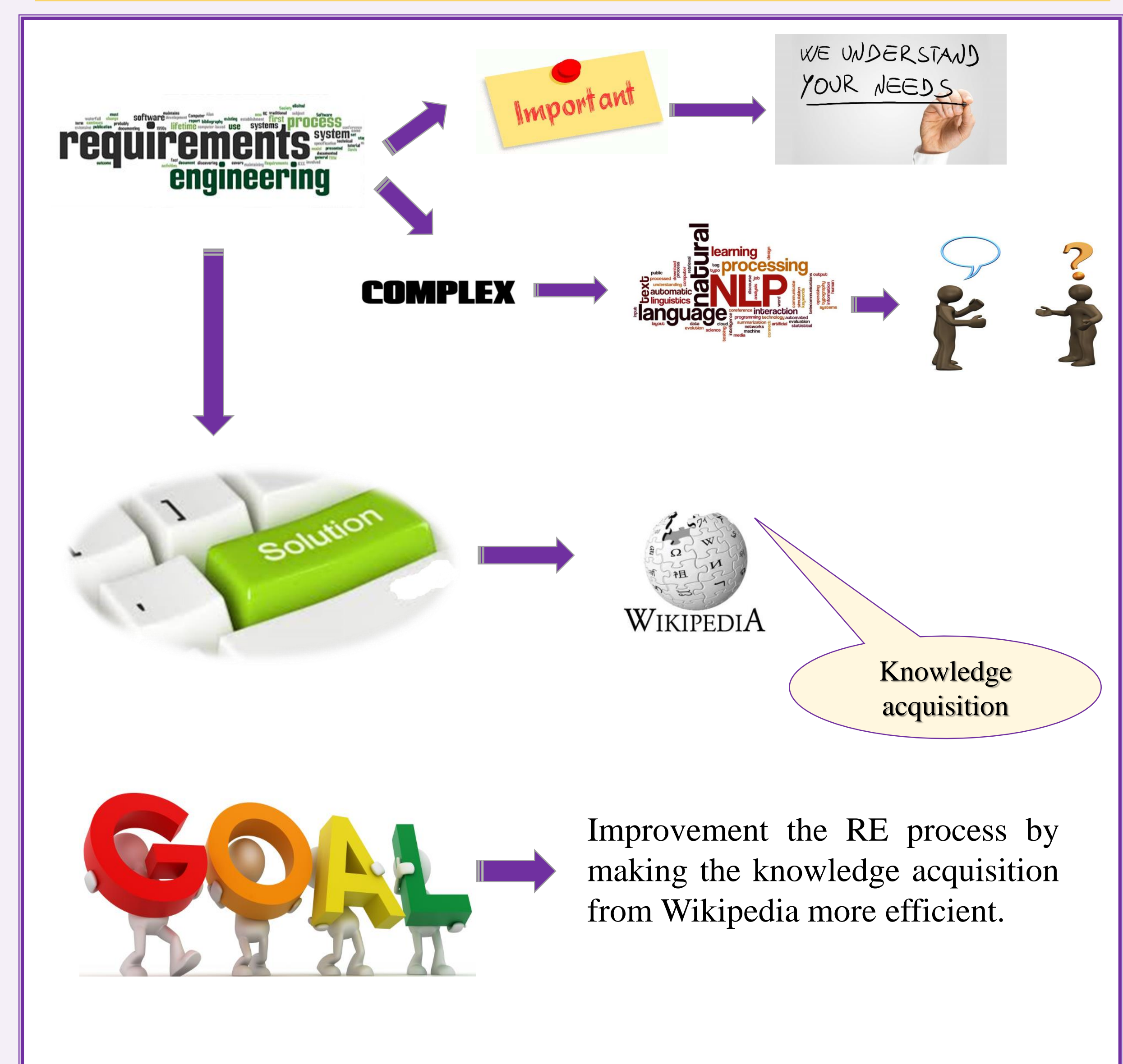
t_1 t_2 ... t_i t_{i+1} ... t_{i+m-1} ... t_n

shift

p_1 p_2 ... p_m

RESULTS

Winner → REGULAR EXPESSIONS performed better in every test.

TECHNIQUE PERFORMANCE

— REGULAR EXPRESSION — STRING SEARCH ALGORITHM

NUMBER OF ARTICLES	REGULAR EXPRESSION	STRING SEARCH ALGORITHM
0	0	0
100	77	249
1000	794	1,279
3000	1,961	2,497
5000	3,129	3,761
10000	5,698	6,853
15000	8,325	9,756
30000	11,137	11,992
50000	22,546	24,346
100000	37,602	41,785
300000	92,333	94,731
500000	140,326	141,070
700000	182,186	182,613
1000000	242,226	260,065

CONCLUSIONS

requirements engineering

Important → WE UNDERSTAND YOUR NEEDS

COMPLEX → NLP natural language processing

Solution → WIKIPEDIA → Knowledge acquisition

GOAL → Improvement the RE process by making the knowledge acquisition from Wikipedia more efficient.

REFERENCES

[1] Mahmoud, A., & Carver, D. (2015, August). Exploiting online human knowledge in Requirements Engineering. In Requirements Engineering Conference (RE), 2015 IEEE 23rd International (pp. 262-267). IEEE.

[2] Hoss, A. M., & Carver, D. L. (2007). Towards combining ontologies and model weaving for the evolution of requirements models. In Innovations for requirement analysis. From stakeholders' needs to formal designs (pp. 85-102). Springer Berlin Heidelberg.

[3] Wagner, C. (2006). Breaking the knowledge acquisition bottleneck through conversational knowledge management. Information Resources Management Journal, 19(1), 70.