



INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

**Automated Generation of Information Architecture from User Stories using
Natural Language Processing**

A Dissertation by

Ms Anjalika De Silva

Supervised by

Mr John Sriskandarajah

Submitted in partial fulfilment of the requirements for the BEng in Software Engineering
degree at the University of Westminster.

May 2022

Abstract

Precise requirement understanding and meticulous formulation of Information Architecture (IA) undoubtedly build the foundation for successful project realisation. In a world exploding with a plethora of information, IA is considered the bearing pillar of a good User Experience (UX) in any platform. However, to achieve the desired level of IA sophistication, a crystal-clear understanding of natural language requirements is crucial. Unfortunately, natural language comes with the baggage of ambiguity. An attempt to map these two disciplines brought to light the research field of Natural Language Processing (NLP) on user stories. An extensive Literature Review (LR) on the domain unveiled the existence of a distinct research gap (discussed in [section 1.4](#)), which ignited this study.

This research project proposes an experimental approach to identify the best-fit mechanism to extract implicit information from user stories after analysing its semantic structure to derive page titles of a system using NLP. A series of experiments were conducted to identify the most promising mechanism to solve the identified problem, from which Deep Learning (DL) showed potential in reaching the desired level of performance.

The results exemplified promising results for implicit information extraction from user stories to derive page titles of a system. Nevertheless, there is space for the solution to improve with the availability of a larger dataset. Furthermore, it paved the way to explore NLP on user stories research with semantic approaches and expand the field with DL.

Keywords: Natural Language Processing, Deep Learning, Information Architecture, Human-Computer Interaction, User Experience, User Interface

Subject Descriptor:

- Computing methodologies → Artificial intelligence → Natural language processing → Information extraction
- Human-centered computing → Interaction design → Interaction design process and methods → User interface design