



Exposé für eine Bachelorarbeit zum Thema:

**Erstellen einer standardisierten Material-Kostengliederung für
Projekte einer Bausoftware mittels Natural Language Processing**

Florian Weidner
Fakultät für Informatik

WS 2022/23

Inhaltsverzeichnis

1	Motivation	3
1.1	Problemstellung	3
1.2	Ziel	3
2	Ausgangslage	4
3	Vorläufige Gliederung	4
4	Zeitplan	7

1 Motivation

1.1 Problemstellung

Das übergeordnete Ziel für die Bausoftware ORCA AVA ist es, möglichst viele Informationen aus 3D-Modellen von Gebäuden in der Ausschreibungssoftware automatisch importieren zu können. Das spart dem Architekten am Anfang in der Ausschreibung viel Zeit. Der für die Bachelorarbeit relevante Teil ist die Übernahme der Baustoffe und Materialien des Objektes/Bauteils. Aus dem Modell übernommene Bauteile sollen also automatisch einer im vor hinein generierten Material-Kostengliederung zugewiesen werden. Als öffentliches Standardformat für 3D-Gebäudemodelle gibt es IFC, welches auch in der ORCA AVA benutzt wird. In diesen Modellen können auch die Materialien der einzelnen Bauteile spezifiziert werden. Diese können an verschiedenen Stellen an einem Bauteil im Modell angegeben werden. Für Materialbezeichnungen gibt es auch keine richtigen Standards. Hier besteht das Problem, dass das Textfeldern für die Materialangabe ein offenes Textfeld ist und somit keine feste Vorgabe existiert und jedes IFC-Modell anders strukturiert ist. Deswegen ist das Ergebnis jeder Kostengliederung unterschiedlich und vom jeweiligen Modell abhängig.

1.2 Ziel

Ziel ist es eine Kostengliederungsstruktur in der Bausoftware ORCA AVA aus den Materialeininformationen einer IFC-Datei zu generieren. Diese Kostengliederung kann am Anfang eines Projektes einmal importiert werden. Wenn dann im Laufe des Ausschreibungsprozesses ein Bauteil aus der IFC-Datei in die ORCA AVA übernommen wird, soll es automatisch einem Material in dieser Kostengliederungsstruktur zugewiesen werden. So kann man das ausgeschriebene Gebäude nach den Materialangaben sortieren und auswerten.

Ein Algorithmus soll zuerst die Möglichkeiten der Materialangabe in der IFC-Datei zusammenführen. Mit Hilfe von Natural Language Processing und Artificial Intelligence eine Lösung entwickelt werden, um einen klassifizierten und standardisierten Materialbaum zu erschaffen. Dieser Baum muss am Ende in das schon von ORCA vorhandene Format einer Kostengliederung abgebildet werden.

2 Ausgangslage

Das Projekt basiert vor allem auf das Themengebiet von Natural Language Processing (NLP). Mit Algorithmen wie WordNet oder Textklassifizierungen.

Innerhalb der Firma gibt es noch fast kein Wissen über dieses Themengebiet. Das Projekt ist also hier ein Test um Themen wie maschinelles Lernen oder NLP in die Entwicklung zu integrieren.

Für das Auslesen von IFC-Dateien wird das Open Source Framework XBim verwendet. Über diese Schnittstelle können die einzelnen Positionen der möglichen Materialangabe ausgelesen und dann zusammengeführt werden. Auf der anderen Seite sind die Abbildungen einer Kostengliederung schon vorhanden. Diese ist eine simple Baumstruktur, die die Oberfläche anzeigen und anpassen kann. Technisch ist gewünscht, dass das Projekt in CSharp zu implementieren ist. Es dürfen zusätzlich auch Python-Skripte verwendet werden.

3 Vorläufige Gliederung

Siehe Seite 5 und 6.

Inhaltsverzeichnis

1	Einleitung	4
1.1	Ausgangssituation	5
1.2	Motivation	5
1.3	Methodik	5
1.3.1	Wissenschaftliche Vorgehensweise	5
1.3.2	Zusatzinformationen Quellen	5
2	Grundlagen	5
2.1	Projektmanagement	5
2.1.1	Vorgehensmodell	5
2.1.2	DevOps	5
2.2	IFC Format	5
2.2.1	Geschichte	5
2.2.2	Dateiformat	5
2.2.3	Verwendung	5
2.2.4	Baustoffe in IFC-Dateien	5
2.3	Möglichkeiten für die Materialangabe eines Bauteils	5
2.4	Das Format Kostengliederung in der ORCA AVA	5
3	Problemstellung und Anforderungen	5
3.1	Problemstellung	5
3.2	Anforderungen	5
3.2.1	Funktionale Anforderungen	5
3.2.2	Weitere Anforderungen	5
3.2.3	Ziele	5
4	Theoretische Konzeption für die Erstellung der Material-Kostengliederung	5
4.1	Textklassifizierungsalgorithmus	5
4.2	WordNet	5
4.3	5
5	Gegenüberstellung der möglichen Konzepte	5
5.1	Messkriterien	5
5.2	Vergleich der Konzepte	5
5.3	Festsetzen eines Algorithmus	5
6	Praktische Umsetzung	5
6.1	Zusammenfassen der Materialschnittstelle einer IFC Datei	5
6.1.1	Entwurf des Algorithmus	5
6.1.2	Implementieren des Algorithmus	5
6.2	Standardisierung der Materialnamen	5
6.2.1	Nutzen von Artificial Intelligence	5
6.2.2	Erstellen einer Datengrundlage	5
6.2.3	Implementierung der Standardisierung	5

6.3	Erstellen einer Kostengliederung	5
6.3.1	Implementieren	5
7	Maßnahmen zur Qualitätssicherung	5
7.1	Clean Code	5
7.2	Technische Hilfsmittel	5
7.3	Tests und Abnahme	5
8	Abschluss	5
8.1	Bewertung der praktischen Umsetzung	5
8.2	Fazit	5
8.3	Ausblick	5
A	Erster Abschnitt des Anhangs	6

4 Zeitplan

- 1.** Recherchieren über verschiedene mögliche Techniken und Algorithmen (2 Wochen)
- 2.** Austesten der gefundenen Algorithmen (2 Wochen)
- 3.** Messen der Ergebnisse der Algorithmen (1 Wochen)
- 4.** Implementieren des besten Ergebnisses (1 Wochen)
- 5.** Schreiben der Arbeit (ca. 6 Wochen)