

Verteidigung der Bachelorarbeit

“Messung der Informationstypen -
Häufigkeiten in der Python-
Dokumentation”

Betreuer: Lutz Prechelt

Sven Wildermann
23. Oktober 2014

Freie Universität



Berlin

Motivation

Verständnis

Verbesserung

Erfolg

Analyse

Vorausgegangene Studie

Walid Maalej and Martin P. Robillard

über

Patterns of Knowledge in API Reference
Documentation

Die 12 Informationstypen

1. Functionality and Behavior
2. Concepts
3. Directives
4. Purpose and Rationale
5. Quality Attributes/ Internal Aspects
6. Control-Flow
7. Structure
8. Patterns
9. Code Examples
10. Environment
11. References
12. Non-Information

`public void addItem(String item)` **A**

Obsolete as of Java 2 platform v1.1. Please use the add method instead.

Adds an item to this Choice menu.

Parameters:
item - the item to be added

Throws:
[NullPointerException](#) - if the i

Type	Method	B
Package	java.awt	
Parent class	public class Choice	

Functionality and Behavior Purpose and Rationale Structure Environment

Concepts Quality Attributes and Internal Aspects Patterns External References

Directives Control-Flow Code Examples Non-information

C

Describes quality attributes of the API, also known as non-functional requirements, for example, the performance or security implications of using the API element (or related elements), such as resources consumed or the execution time of a method. Also code with a value of true if the block provides information about API's internal implementation that is only indirectly related to its observable behavior. For example, indicates the main data structures and algorithms employed.

- This is a "graceful" release and may cause IO operations for consuming the remainder of a response entity.
- Enumerating through a collection is intrinsically not a thread-safe procedure.

D

Save and Finish Next >

CADo – Content Analysis for Software Documentation

<http://cado.informatik.uni-hamburg.de/tool/>

Struktur der Elemente

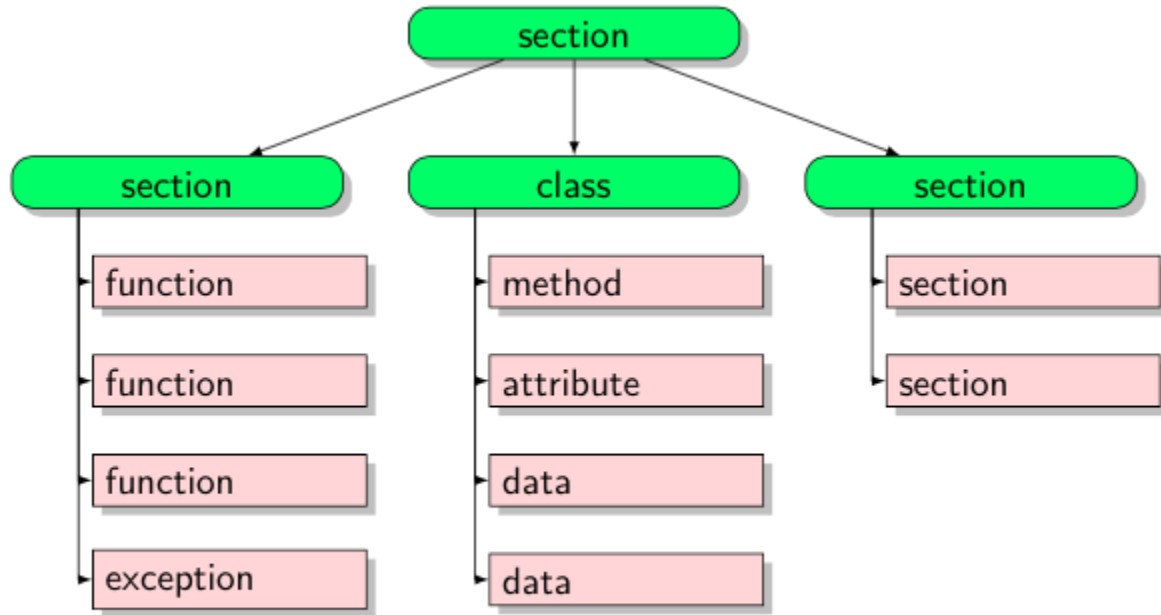


Abbildung 10: Beispielhafte Struktur der HTML-Elemente

Einteilung in Kategorien

1. Methoden

method, classmethod, staticmethod, function

2. Felder

attribute, data

3. Module

section

4. Klassen

class, exception

5. Beschreibungen

describe

Zwischenergebnisse

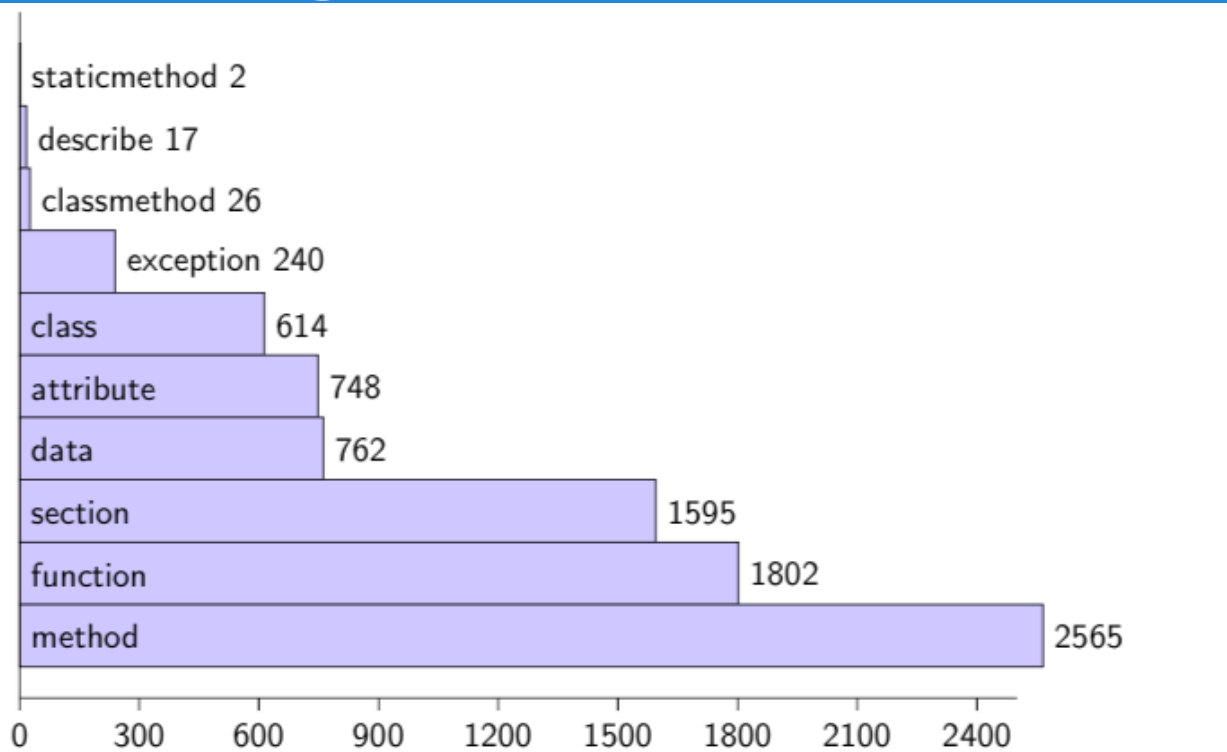
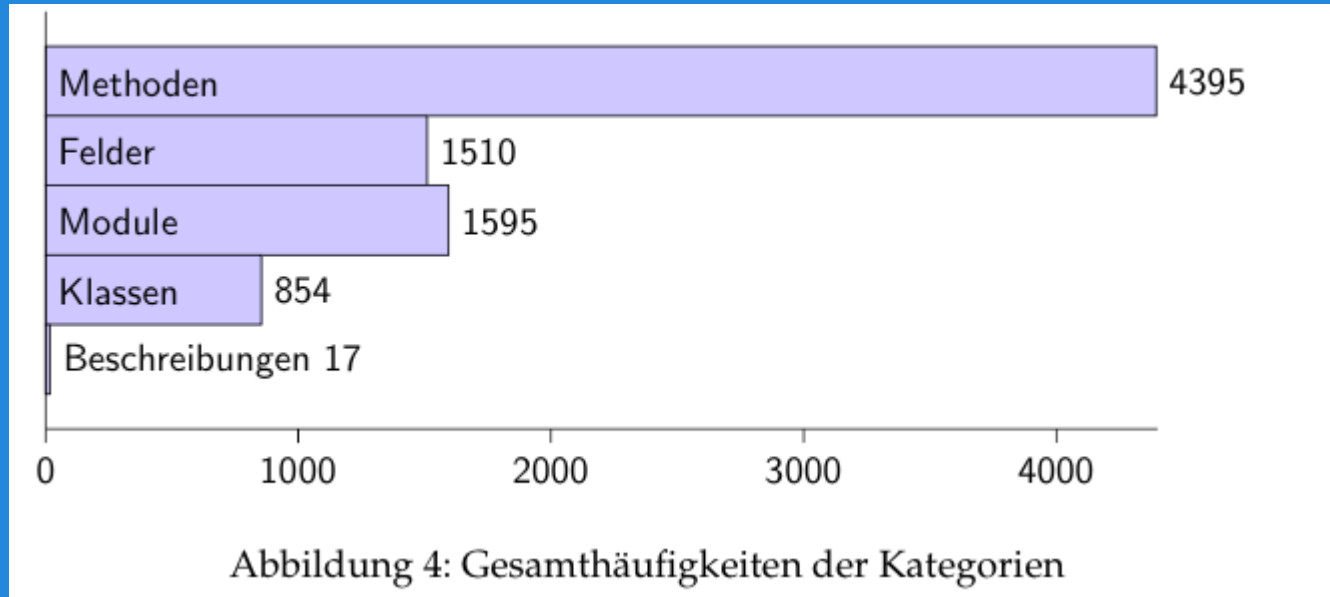
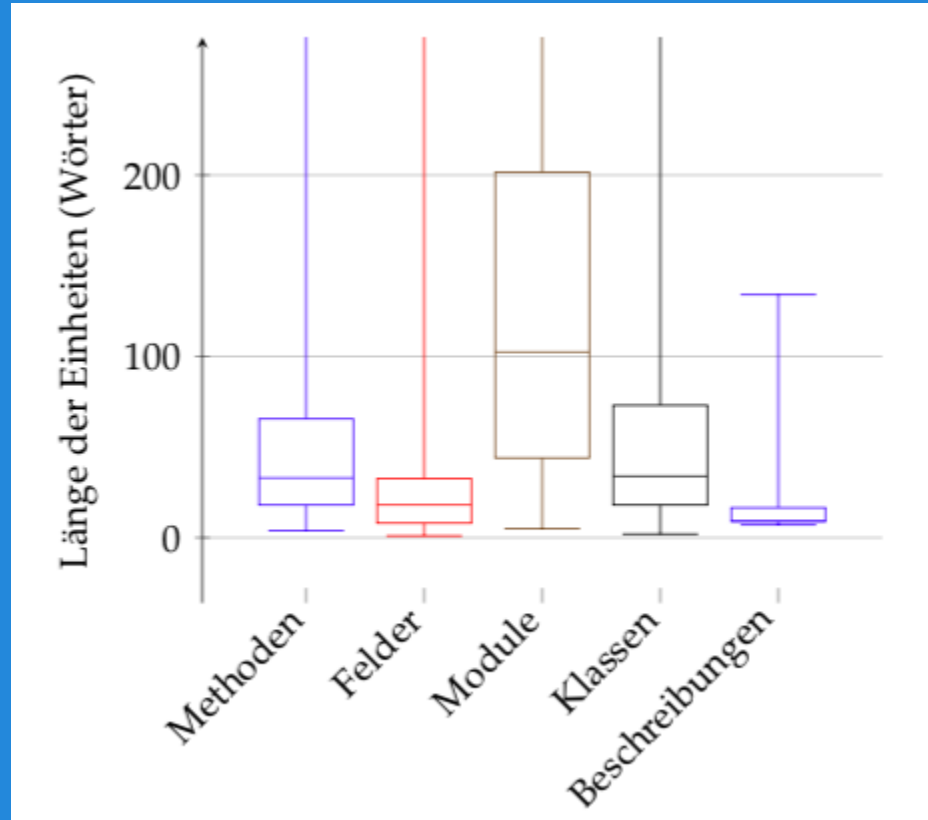


Abbildung 3: Gesamthäufigkeiten der Einheiten

Zwischenergebnisse



Zwischenergebnisse



Methodische Abweichungen

- Definitionen der Informationstypen
- Markierung von Informationstypen

Live Demonstration

Übereinstimmung zw. Gutachtern

Markierung 1



Markierung 2

min. 50% Überlappung

der Länge beider Markierungen

Gutachterübereinstimmungen

	Studie von Maalej und Robillard	diese Bachelorarbeit
kleinster Min-Wert	0,309	0,703
größter Max-Wert	0,981	1,0
Median	0,820	0,879

Ergebnisse

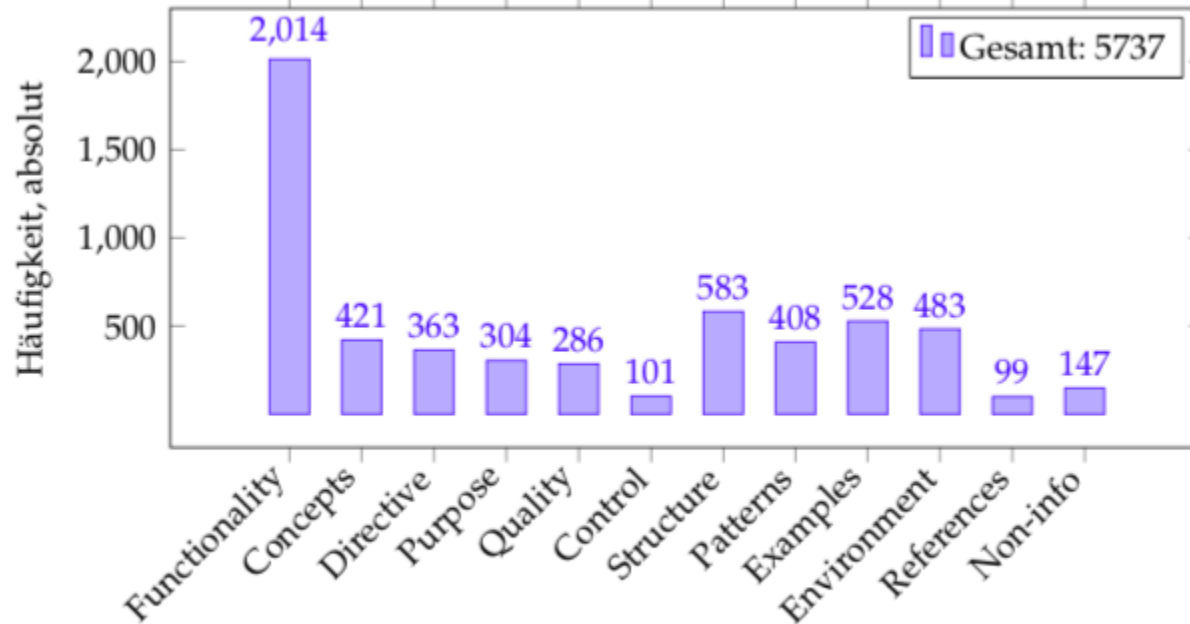
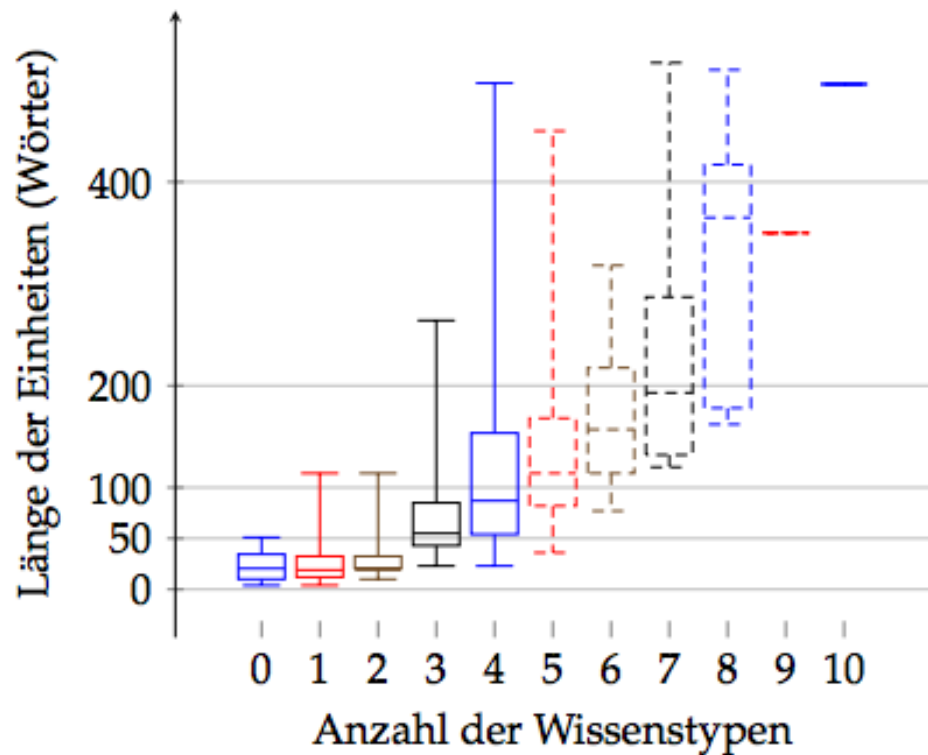


Abbildung 28: Häufigkeiten der Wissenstypen

Ergebnisse



Zusammenfassung

- verbessertes Verfahren
bzgl. Sorgfalt und Detailreichtum
- weniger “Nicht-Information” in Python
- #Wörter ~ # Wissenstypen

Verteidigung der Bachelorarbeit

“Messung der Informationstypen -
Häufigkeiten in der Python-
Dokumentation”

Betreuer: Lutz Prechelt

Sven Wildermann
23. Oktober 2014

Freie Universität



Berlin