# Taxonomie von SE Domaenen / Kontexten

## Protokoll Zwischenpräsentation 0X.02.2021

Weiteres Vorgehen (F\_)

Threats (F\_)

* In wie weit muss ich Maßnahmen treffen um diese zu mildern
* In wie weit kann ich diese wegdiskutieren

## Protokoll Zwischenpräsentation 28.01.2021 LP Fazit

Fokus: / Kern!!

1. **Definition: Was ist Kontext? Was ist ein Kontextfaktor?**
2. **Motivation und Sammlung: Welche Sorten von Kontextfaktor bringen aus welchem Grund in welchen Fällen einen großen Nutzen für wen?**

Vorgehen :

Suche nach einer Handvoll von Beispielen für Punkt 2.

> Kern der Taxonomie, Antwort für Punkt 1

Falle von a) Quantifizierung oder b) zu viel Programmierung. Das liegt beides nicht im Kern, aber wenn Sie den Kern nicht hinbekommen, wird das Ergebnis zwangsläufig schlecht.

## Protokoll Zwischenpräsentation 28.01.2021 Notizen

LP Notizen

Ziele (F5)

* Müssen evaluiert werden
* 2.Ziel eher erhoffter Nutzen als Ziel
* 3.Ziel Mittel zum Zweck

Hintergrund Kontexte (F7)

Bessere Definition ¡! Zu sehr zweckbezogen

LP Vorschlag: Ein Kontext defniiert den Bereich über den die Ergebnisse sprechen

Vorgehen Sichtung (F13)

* Kritik: Badge sagt mehr über Art der Daten als Qualität und Sorgfalt aus
  + Paper mit vertraulichen Daten (keine Change auf Badge Artifact Available/Reusable)
  + Paper mit manueller Analyse: Repositories, Commit Daten haben es einfacher
  + Badge-Entscheidung fragwürdig, Dokumentieren und in Frage stellen, als Verteidigen
* Evaluation der Annahme auf F13 nötig, keine Verteidigung, sondern Selbstreflektion/Kritik
  + Schlechte Entscheidungen nicht verteigen
  + Schreiben: dass man erhofft hat die Paper mit mehr Qualität zu erwischen, aber nicht weiß/nachweisen kann, ob es geklappt hat, wegen denn und den Gründen

Vorgehen Charakterisierung (F19)

* Keine Nachkommazahlen, aufgerundet auf ganze Projektzahlen oder Anzahl der Paper (bei niedgrigeren Zahlen)
  + Nachkomma zu abstrakt und ablenkend

Vorgehen Analyse zur Verallgemeinbarkeit (F23)

* Bitte semantisch zusammenfassen: Limitations ist ca. Threats to Validity
* Quote 2 / Quotes in dieser Klasse
  + Ausreden aufspüuren, Draufhauen
  + 12 Applications können nie represantativ für alle realworld Applikationen sein
  + Seiten gegeneinander auspielen: Linke Seite may not represent Rechte Seite
* Immer Gesamtheit des Papers betrachten
* Verallgemeinerbarkeit !!!
  + TODO in Präsi / Tool / Masterarbeit überall korregieren

Kontexte (F31)

* Einstufungen
  + Allgemeinwissen, leicht erkennbar (evtl schon vom Titel ableitbar)
  + Wenn man sich im Thema auskennt, erkennt man Kontexte
  + AutorInnen müssen es explizit sagen, sonst erkennt niemand den Kontext
* Evaluieren, so wie LP es gemacht hat. Lassen sich die identifizierten Kontexte aus den Titel ableiten > Sollte es ein Qualitätskriterium geben, dass man mittels Titel den Kontext einer Arbeit erkennen kann und in die Landkarte einordnen kann
* Sourceing: Quellen der Daten mitdokumentieren

Kontext Definition > Tool, Zahlen, Statistiken

Konkrete Beispiele > Tool, Zahlen

* Negativ Beispiele Paper, die sich nicht leicht einordnen Lassen in Taxonomie
  + Für Gummipunkte von Franzis Zierris !

Konkreter Ansatz (<-> abstrakter Ansatz = schwieriger)

* Sensemaking mittels Beispiel Exemplare
* Paper finden, wo es nützlicher wäre eine Taxonomie zur Hand zu haben für Zweck / Use Case XY – je vers desto besser
* Begründung von Dimension (F31) durch Beispiel Exemplare – Argumentation, warum diese Dimension relevant ist

Rückblick warum man diese Arbeit macht?

* Grundlage für Vermutung von Verallgemeinbarkeit, dass AutorInnen in Zukunft besser Verallgemeinbarkeitsaussagen treffen
* Lernen ähnliche Kontexte oder nicht ähnliche zu erkennen > Landkarte
* Terminologie > strukturiert reden & reden können

Relevanz, Nutzen der Taxonomie ausarbeiten

Sonstige Sachen zum Schreiben:

* Evaluieren
  + Ist es eine offene Liste / 1.Schritte zur Taxonomie oder abgeschlossen Liste
* Evaluation des mögl Nutzen einer Taxonomie / Beweis in Future Work
  + Exploratory Untersuchung, wie es im Prozess/ Vorgehen der Wissenschaftler nützlich sein kann
  + Rolle:
    - Autor
    - Leser
    - ICSE Konferenz Gutachter
    - Studis – (die an Seminare teilnehmen, Einordnung der Literatur, Auswahl der Paper für SE Beiträge, sodass Paper Kombination nicht zu nah beieinander sind)
* Analyse: Kontexte – Vorkommen

Frage aus Analyse Verallgemeinbarkeit: Warum mögen Autoren sich nicht spezialisieren? Festlegen?

* UseCase: Taxonomie / Landkarte als Hilfe, Guide, um sich festzulegen
* Können Fragen die bei den Analysen auftauchen als UseCase betrachtet werden?

Feedback zum Vortrag:

* Dass man in die Folien reinspringen kann von Slide X zu Slide Y ohne den Faden zu verlieren
* Viel Text aber mit Sinn und Zweck
* Content: Weniger über bisheriges Vorgehen, weil Ziel des Vortrags: Feedback, Diskussion
* Amerikanischer Vortragsstil – Storytelling: Anfang bis Ende (Vorgehen, Ziele)
  + Macht ein andere Stil mehr Sinn für die Masterverteidigung?

Für die Masterverteidigung

* Präsentation der Leistung, Ergebnisse und Nutzen, nicht des Vorgehens
* Ziel des Vortrags: Gummipunkte zu erhalten, gute Note

## Protokoll Sprechstunde 14.01.2021

Mit Lutz Prechelt (LP)

Annahme: Ideen in SE nicht universell, sondern bewegen sich in einem Kontext, Dimension

Thesis: Versuch datengetrieben die SE Welt zu definieren

Mein Annahme Sheet: Vorstellung von Beschaffenheit / Qualität der ICSE (Schwächen)

Paper Charakterisierung: Prozess vs Produktfragen

Paper mehr in der SE-Welt als in der Welt des Papers betrachten

* Zweck: Anwendung, Verallgemeinbarkeit

Bei zu wenig Kapazität nicht alle Paper lesen

* Auswahl per Kriterium, begründet, vielversprechend Effizienz <-
* Vs Zufall

Mehr Zeit in sensemaking als Analyse des Rohmaterials

28.01 Zwischenvortrag

- Folien nur mit spezifischen Fragen schicken

## Protokoll Sprechstunde 14.01.2021

Mit Franz Zieris (FZ)

Kriterium

* Explizite Aussagen zur Verallgemeinbarkeit
* Explizite Aussagen zur Grenze der Anwendbarkeit der eigenen Ergebnisse

Knowledge seeking vs solution seeking Paper

Treats an meiner Forschungsarbeit

* Keine Forschungserfahrung
* Keine Vorkenntnisse vom Material

Mein Analyse Tool

* Motivation: Grenzen von anderen Tools wie MaxQDA
* Als Werkzeug für die Thesis, nicht als Ergebnis, um Zsmhänge zu erkennen

Titeländerung: 10 häufigsten Probleme bei ICSE Paper zur Identifikation von SE-Kontexten

**Schritt 1**: Papers mit guter Diskussion, klaren Aussagen und stimmigen Annahmen finden bzw expliziten Aussagen zur Verallgemeinbarkeit oder zur Grenze der Anwendbarkeit der eigenen Ergebnisse

* Diese Paper als Arbeitshypothese (Vorbilder?) nutzen

**Schritt 2:** Aus diesen Papers rausziehen, welche Dimensionen/Kontexte es gibt

* zB. Kritische Anwendung (Raumfahrtforschung) vs nicht kritische Anwendungen (mobile Apps)

**Schritt 3:** andere Paper in diesen Dimensionen einordnen, ob sie links oder rechts liegen

**Schritt 4:** aus den Dimensionen eine Taxonomie erstellen, die als ‘Landkarte’ genutzt werden kann, um zukünftige ICSE Paper in ihren Kontexten einzuordnen

Die Annahme hinter dem Vorgehen ist: „Gute SE-Forscher haben eine Proto-Taxonomie bereits im Kopf und benutzen diese in ihren Diskussionen.“

Hier ein paar Ergänzungen:

zu Schritt 1: “Vorbilder” in der Hinsicht, dass die Autor/inn/en sich schon gute Gedanken gemacht haben, und man deren SE-Taxonomie(-Vorform) benutzen kann.

zu Schritt 3: Das Einordnen ist nicht die Hauptsache, sondern das Prüfen, welche Dimensionen etwas taugen, also zu einer sinnvollen Unterteilung der Forschungslandschaft bzw. des Software Engineering führen.

Schritt 3a: Nachschärfen, was die Dimension eigentlich ist. Es sollte im Laufe der Zeit immer klarer und einfacher werden, neue Paper einzusortieren.

Schritt 3b: Paper, auf die sich mehrere Dimensionen nicht sinnvoll anwenden lassen (oder sogar keine), kommen auf einen Sonderstapel.

Schritt 4a: Gesamtheit des Sonderstapels betrachten: Was ist an denen besonders? Sind die in einem Teil der SE-Welt unterwegs, der noch keinen Namen hat (= weitere Dimension)? Oder sind die einfach komisch und anders?

## Protokoll Sprechstunde 17.12.2020

Qualität der Forschung verbessern

* Untersuchung, es gilt X / Annahmen unter welche x gilt benennen
* Stillschweigende Annahmen, dass immer über X Software geredet wird, aber auch in Software Y interessant ist
* Selbstverständliche Annahmen
* Verallgemeinbarkeit diskutiert innerhalb der Welt von Software X, die nicht besprochen wird
* Annahmen > construct Validität, interne Validität, Interpretation, Interpretationschritte bei den Vorgehensschritten (!! -> 2.Iteration)
* Ergebnis fragwürdig wenn Annahmen nicht stimmen

Keine Stellungnahme zur Verallgemeinbarkeit -Warum?

* Keine Relevanz
* Kein Fokus?
* Selbstverständlichkeit

Indizen/Zitate sammeln um selbst Annahmen aufzustellen

Stellen: TtV, Limitations, FW gut für bewusste Annahmen, zu sehr an der Oberfläche

Systematisches Suchen/Finden von Annahmen -> Stellen: Vorgehen, Interpretation, construct, interne validität

Analyse der Annahmen

Klare, simple, bewusste Annahmen vs Annahmen die schief gehen könnten

Vs unbewusste, implizite Annahmen

Mehr Mails !!

## Protokoll Sprechstunde 20.11.2020

Motivation der Thesisseite / Sprechstundennotizen mehr auf Anwendbarkeit / Verallgemeinbarkeit gerichtet

* Aber in Related Work (durch Eingrenzung, Abgrenzung) Kontextfaktoren klarer beschrieben
  1. Themenkreise, implizites Verständnis von Verwandtschaften
  2. Wichtigkeit von anderen Arbeiten ist unklar und unterschiedlich in Studien, auch wenig Aufbau auf andere Arbeiten als Grundlage

Kontexte

* spannen von Dimensionen
  1. Technologische Sphären
  2. Kleine, medium, große Teams
* Konkretheit von Annahmen
* Einzelne Attribute von Verwandtschaften

Subjektives Fingersaugen

Vollständigkeit kein Problem

* Übersehen nicht schlimm
* Sind meine Kontextfaktoren wichtig?
* Textschnitt -> Betrachtungsweise/Argumentation -> Dimension
  + Argumentationsstrang: Rohdaten > Schlussfolgerung

2.QS durch Lesers eigene Meinung, weil Evidenz offenlegen an Leser

Begruendungsbedarf, warum Arbeit wertvoll / Aufwand

Text:

Teilergebnisse/Analyse/Argumentationsstrang

+ Ausblick + Einschätzung von Ergebnis / Anzahl von diesen Argumentationssträngen

Ergebnisse

* Taxonomie
* Arbeitsweise, Lesetechnik, Wann sind welche Paragraphen interessant
* Denkmuster für implizite/versteckte Annahmen !

## Protokoll Sprechstunde 22.10.2020

Software zur qualitativen Datenanalyse: MaxQDA

* Notiz/Memo an Textschnipsel
  + Was verstehe ich unter Konzept, Merkmale
  + Sonstige Insights
* Labels für Konzepte an Textschnipsel

Methode des ständigen Vergleichens

GTM -> TODO einlesen

TODO aufräumen

Immer Reihenfolge, die am meisten Erkenntnis bringt

* Vorwärts bis Erknenntnis, dann rückwärts: bereits gelesene Paper anschauen und Erknenntnis evaluieren –
* 3,4 gewinnbringende Kenntnisse in MA beispielhaft dokumentieren um Arbeitsweise zu dokumentieren, ansonsten Strichliste

Nicht in Trott fallen Lassen und immer gleiche Konzepte ohne Erkenntnisse anwenden

* Theoretische Sättigung?
* Immer hinterfragen!

## Protokoll Sprechstunde 24.09.2020

Reihenfolge anhand der Awards und roten/grünen Siegeln

* 1. Kontextfaktoren gefunden
* Expliziter als die zufällig ausgewählten durchgelesenen Arbeiten

Wenn hinterher zu viel Kapazität über, oder keine Ergebnisse zur Taxonomie

* Untersuchen, warum keine Kontexte in den Arbeiten vorhanden sind, kein Anspruch auf Universalität ?
* Interviews von Prof. Dr. Prechelt auf der ICSE 201X analysieren
  + Audiodateien vorhanden, “Semi”-Transkription notwendig

Anmeldung

* Als Startschuss, Sicherheit, dass Ergebnisse geliefert werden können

## Protokoll Sprechstunde 03.09.2020

Implizit, selbstverständlich für Autorinnen

Reihenfolge kann zum Übersehen führen, Effekt

Reihenfolge nach Effizienz wählen, damit Licht früh aufgeht, danach immer Review

* Heterogen besser, Arten von Artikel, Sorte von RQ

ScanLeseMethode graduell entwickeln

Ordnung, Gruppierung von Kontextfaktoren, Beziehungen dazwischen

Warum sind welche Faktoren leichter sichtbar und welche nicht, Natur der Artikel?

Empirische Angehensweise:

Alle Artikel durchgehen

<->

Theoretische Rangehensweise:

Hypothesen erstellen und Beispiel Papers finden

Idee:

-> Reihenfolge Paper mit Qualitätssiegel

-> Danach Charakterisieren aller Arbeiten

-> und heterogene Reihenfolge nach Artikel Typ

## Protokoll Sprechstunde 06.08.2020

Kontextfaktoren = in welchen Kontext lassen sich Ergebnisse eines Papers verallgemeinern? Grenzen von Verallgemeinbarkeit

In welchen Kontexten wurden Daten erhoben? Ränder von Kontexten?

Formulieren Autoren Einschränkungen? Anflüge von Kontextfaktoren, implizite Aussagen. Wie reden Autoren über Kontext

🡪 am Ende schöne Fallbeispiele

Threat of Validity, sollte viel drin stehen, aber oft nur Blabla

TODO

Begriffe hinterfragen, zb. Commits, LoC – Gewicht?

Öfteres Kodieren, semantische Konzepte sammeln

Klassifizieren von Kontext

## Protokoll Sprechstunde 09.07.2020

Bisher: grobe Unterscheidungen zwischen Projekten/Kontexten in Studien/Artikeln

* Informationssystem vs Echtzeit, OS, App (low critical systems) vs high critical sytems

ETA**T** Ziele (Four presumed gaps in the SE research community’s knowledge)

* Taxonomie erstellen ueber verschiedene Software Kontexte

Faktoren ableiten, die Softwareprojekte als Tuple klassifizieren

* zB. Team Groesse, eingesetzte Technologien, ...
* Sind alle gleichzeitig relevant?
* Wichtige Kombinationen?

Literaturstudie (da Feldstudie zu aufwaendig)

* ISCE 2020, 129 Artikel, Technical Research Papers analysieren
* Was fuer Faktoren wurden genutzt zur Definition des Problembereichs, zur Auswahl von Projekten -> Kontextfaktoren
* Welche Rollen spielen verschiedene Faktoren? Welchen Zweck erfuellen Faktoren?
* Klassen von Softwareprojekten/Kontexten aus Faktoren bilden -> Tuple
  + Wie viele? Wie Abgrenzung und warum?

🡪 sensemaking

ISCE 2020 Artikel

* Hohe Qualitaet, aber nicht immer relevant
* Existenzberechtigung des Themas, starke/schwache Relevanzbegruendung?
* Wieso gibt es weniger relevante Themen bei der ISCE?

Grobes Vorgehen

* Artikel ueberfliegen, Stellen ueber Kontextfaktoren finden und markieren
* vereinzelt Artikeln lesen (pro Kategorie)
* Methode entwickeln um Stellen zu finden und Abschaetzen, wie viele Stellen erwischt, wie viel Prozent verpasst (false negatives)
* Software zur qualitativen Datenanalyse zum Sortieren, Statistik, Datenvisualisierung

Erste Schritte

* Artikel Sorten/Typen kategorisieren = Vermutungen
  + Basierend auf Titel und Abstract
* Fuer jede Kategorie, ein Artikel aussuchen und Faktoren suchen