Bachelorarbeit

Schweizer Firmendynamik: Ein Überblick

Haben IT-Firmen eine höhere Austrittswahrscheinlichkeit?

eingereicht an der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Universität Bern

Institut für Wirtschaftsinformatik
Dozentur Digitale Nachhaltigkeit

Dr. Matthias Stürmer

eingereicht von
Dominic Schweizer
von Neckertal, SG
im 07. Semester
Matrikelnummer: 10-808-988

Studienadresse
Dorfberg 555
3550 Langnau i.E
079 819 91 10
domi@hilotec.com

Bern, 31.01.2019

Zusammenfassung

Das Crawling des Schweizerischen Handelsamtsblattes ermöglicht Einsichten in die Firmendynamik der Schweiz im Zeitraum von 2002-2018. Die aus den Meldungen ermittelte Lebenszeiten der Unternehmen zeigen Ähnlichkeiten zu einer Exponentialverteilung. Durch die Applikation von Text Mining Methoden konnten die 408'467 in deutscher Sprache verfassten Neueintragungen in verschieden Kategorien eingeordnet werden. Der Vergleich der Kategorien zeigt, dass deren Überlebensraten unterschiedlich sind. Die Kategorien IT-Unternehmen, Gastronomie, Reinigung haben ähnliche Anteile, jedoch ist die Überlebensrate von IT-Unternehmen mit 55% tiefer als die der anderen Kategorien.

Summary

Crawling the Swiss Official Gazette of Commerce delivers insights in the company dynamics of the swiss economy during 2002 – 2018. The calculated lifetime from the notices shows similarities to an exponential distribution. The author applies text mining methods on the 408’467 entry-notices in German language to label the companies in categories. The categories show different survival rates. The categories IT-Companies, gastronomy and cleaning have similar proportions, but the survival rate of IT-Companies is with 55% lower than the survival rates of the other 2 categories.

Inhaltsverzeichnis

Inhalt

1 Einleitung

Neue Unternehmen schaffen Arbeitsplätze, reduzieren die Arbeitslosigkeit und dienen dem wirtschaftlichen Wachstum. Die Phase nach der Gründung verläuft nicht für jedes Unternehmen gleich. Santarelli und Vivarelli haben in ihrem Paper unterschiedliche Studien zusammengefasst und erläutern, dass eine staatliche Förderung auch unerwünschte Nebeneffekte haben kann. Unternehmen, deren Gründer zu wenig unternehmerische Qualitäten aufweisen oder die Unternehmensgründung als Glückspiel ansehen, haben eine schlechte Performance nach der Gründung. Unternehmen mit einer geringen Performance haben eine geringere Überlebenschance und eine höhere Mortalität. Falls viele neu gegründete Unternehmen eine hohe Mortalität haben führt dies zu Turbulenzen im Sektor. Die Turbulenzen wiederum beeinflussen auch die Unternehmen mit einer guten Performance negativ.[1]

Die beiden Autoren schliessen ihre Untersuchung mit der Einschätzung, dass es vielleicht auch zu viele Neugründungen und zu wenig Unternehmertum gibt.[1]

Coad hat in seiner Studie 2010 die Altersverteilung der Unternehmen untersucht und kam zum Schluss, dass eine Exponentialverteilung geeignet ist, um die Altersverteilung aggregiert und über einen grösseren Zeitraum zu betrachten. Die Studie verwendete dabei Daten der Sozialversicherung. Die Exponentialverteilung ist weniger geeignet um die Verteilung sehr junger oder sehr alter Unternehmen zu modellieren. Ausserdem können einzelne Sektoren andere Verteilungen aufweisen, so dass die Exponentialverteilung vor allem dann geeignet ist, wenn alle Sektoren aggregiert betrachtet werden. [2]

Das Eidgenössische Amt für das Handelsregister genehmigt sämtliche Einträge der kantonalen Handelsregister. Sämtliche grösseren Firmen in der Schweiz müssen im Handelsregister eingetragen sein und das Handelsregister muss sämtliche Einträge im Schweizerischen Handelsamtsblatt (SHAB) veröffentlichen. Welche Rechtseinheiten im Handelsregister eingetragen sein müssen, wird im Obligationenrecht: Handelsregisterrecht geregelt. Wenn eine Firma neu im Handelsregister eingetragen wird, muss im SHAB eine Neueintragung mit Unternehmenszweck und weiteren Informationen veröffentlicht werden. Ändert sie ihren Standort, die Rechtsform usw. oder meldet sie Konkurs an muss eine Mutationsmeldung veröffentlicht werden. Wird ein Unternehmen gelöscht, muss eine Löschungsmeldung getätigt werden.

In der Schweiz existieren Quellen für Open Gouvernement Data [3]. In dieser Bachelor Arbeit wird das SHAB abgefragt, um Einblicke in die Schweizer Firmendynamik zu gewinnen. Sie soll zeigen wie sich die Population der Schweizer Unternehmen im Zeitraum vom 2002 bis 2018 entwickelte. Mit der Beschreibung der Population kann auch Auskunft über die Mortalität von neu gegründeten Schweizer Unternehmen gegeben werden. Neu gegründete Unternehmen werden anhand des Unternehmenszwecks und den NOGA Codes klassifiziert sowie in Cluster aufgeteilt um zu zeigen, dass die Mortalität je nach Sektor variiert.

1.1 Ausgangslage

Seit Coad 2010 in seiner Studie die Lebensdauer von Firmen und deren Verteilung untersucht und das Gebiet noch als unerforscht deklariert hat[2], sind weitere Publikationen entstanden[3], [4], [5], die sich ebenfalls mit dem Phänomen beschäftigen. Keine Publikation beschäftigt sich jedoch mit der Situation in der Schweiz.
Um die Daten für die Untersuchung zu generieren wird in einer Studie das Handelsregister als zuverlässige Quelle dargestellt.[6] Das SHAB bietet sämtliche Publikationen als PDF oder XML zum Download an. Die Meldungen können verwendet werden, um die Lebensdauer und die Mortalität sowie den Zweck der Firma festzustellen und somit den Geschäftsbereich der Firma zu identifizieren.

1.2 Problemstellung

IT-Unternehmen könnten durch ihre geringeren Investitionskosten bei Neugründungen eine tiefere Eintrittsbarriere haben als andere Unternehmen. Damit würden sie eher Revolving-Door-Gründer anziehen, die ein höheres Austrittsrisiko aufweisen. [1] Diese Arbeit soll zeigen wie sich die Schweizer Firmendynamik im Untersuchungszeitraum entwickelt hat. Weiter soll sie die Frage beantworten ob IT-Unternehmen tatsächlich eine höhere Austrittswahrscheinlichkeit haben als andere Unternehmen.

1.3 Zielsetzung

Das Ziel dieser Arbeit ist zu zeigen, dass aufgrund des geringeren Initialaufwandes bei der Gründung von IT-Unternehmen verhältnismässig mehr Austritte als bei den anderen Branchen verzeichnet werden.

Durch eine empirische Vorgehensweise und der Modellierung der Unternehmen anhand der Handelsregisterdaten soll eine vollständige Übersicht der Firmendynamik im IT-Sektor gezeigt werden.

1.4 Aufbau der Arbeit, Methodisches Vorgehen

Um die Fragestellung zu beantworten wurden zuerst Datenquellen für das Handelsregister gesucht und deren Möglichkeiten ausgewertet. Anschliessend wurde eine Übersicht über die Handelsregistermeldungen erstellt. Aus dem Text der relevanten Meldungen wurden Unternehmen modelliert und diese in Cluster aufgeteilt und die Cluster anhand von Wordclouds identifiziert. Im Kapitel zwei werden die theoretischen Grundlagen für die Arbeit gesetzt. Das dritte Kapitel beinhaltet die Datenerhebung über die verschiedenen Datenquellen und den Ablauf der Modellierung sowie die Klassifizierung der Unternehmen. Das vierte Kapitel enthält die Beschreibung der fehlgeschlagenen Versuche. Im fünften Kapitel werden die Ergebnisse erläutert und im sechsten diskutiert sowie die Limitationen der Arbeit erläutert.

2 Theoretische Grundlagen

Die Eintrittsrate der Unternehmen ist definiert als Summe der Unternehmen, die während der Periode neu gegründet wurden. Die Austrittsrate der Unternehmen ist definiert als die Summe der Unternehmen, die während der Periode gelöscht wurden. Coad hat in seinen Untersuchungen festgestellt, dass die sich Eintrittsrate in fast allen Sektoren über die Zeit wenig ändert. Die Ausnahme bildet die Bergbauindustrie, welche eine ausgereifte Industrie ist und somit weniger Eintritte verzeichnet.[2]

Das Alter einer Unternehmung definiert sich in der folgenden Untersuchung als Dauer zwischen der Gründung und Löschung des Unternehmens. Gründungen und Löschungen von Unternehmen werden publiziert, mit der Differenz der beiden Publikationsdaten lässt sich das Alter eines Unternehmens berechnen.

Die Halbwertszeit von Firmen wird in verschiedenen Studien auf drei bis vier Jahre geschätzt.[6]

Coad hat festgestellt, dass die Überlebensrate von sehr jungen Firmen mit dem Alter der Firma wächst und nicht konstant ist. "Die Überlebensrate ist im ersten Jahr am tiefsten und steigt stetig mit der Zeit".[2] Es lässt sich daraus schliessen, dass die Exponentialverteilung für sehr junge Firmen nicht eine passende Verteilung ist, weil die Austrittswahrscheinlichkeit von sehr jungen Firmen höher ist.[2]

3 Methodik

In diesem Kapitel wird der Ablauf der Arbeit von der Datenerhebung über die Modellierung bis zur Klassifizierung erläutert.

3.1 Identifikation der Datenquellen

Wie Coad beschreibt, ist das Handelsregister einer Nation eine wichtige Quelle, um die Firmendynamik zu betrachten.[6] In der Schweiz liegt die Kompetenz für die Handelsregisterführung bei den Kantonen, es gibt jedoch ein Eidgenössisches Amt für das Handelsregister, welches die Tagesmeldungen der Kantone genehmigt. Die Handelsregistermeldungen werden von Montag bis Samstag täglich im schweizerischen Handelsamtsblatt publiziert. Als Datenquellen eignen sich somit das schweizerische Handelsamtsblatt und das Eidgenössische Handelsregisteramt EHRA.

- https://www.shab.ch
- https://ehra.fenceit.ch/de/shab

Für beide Quellen wurden Crawler geschrieben, um ein Datensample zu crawlen und die Datenqualität zu evaluieren.

Neben den Handelsregistermeldungen bietet das EHRA einen Firmenindex mit Suchfunktion an. Im Firmenindex werden die Handelsregistermeldungen aggregiert dargestellt, um einem Nutzer den aktuellen Zustand der Firma anzuzeigen.

- https://www.zefix.admin.ch

Die Suchmaske von zefix.ch erlaubt Suchabfragen nach Firmennamen, CHID und UID. Ausserdem können alle Meldungen innerhalb einer Periode auf die Unternehmen aggregiert werden. Jedoch bietet die Funktion nur den Zeitraum von 02.03.2016 bis zum aktuellen Datum, was nicht dem geplanten Untersuchungszeitraum entspricht.

Eine Untersuchung mit Web Analytics einer Suchabfrage zeigt, dass die Suchmaske die Abfrage als POST Request weitersendet. Ein Versuch zeigte, dass die API offen ist und auf einen Request mit einem JSON-Body und den nötigen Daten antwortet.

Neben den Suchabfragen bietet das EHRA auch noch Jahresstatistiken an, diese werden verwendet um die Güte der Crawler und Parser zu bewerten.

3.2 Erhebung der Meldungen

Die Webseite des schweizerischen Handelsamtsblattes bietet ein Archiv der Meldungen vom Zeitraum 01.16.2002 bis zum 31.08.2018 an. Auf shab.ch/#!/search/archive

können über HTTP GET Request Suchanfragen nach Stichworten und Zeiträume getätigt werden. Die Resultate werden von shab.ch als JSON übermittelt. Das JSON besteht aus einem Array das Metadaten zu den Meldungen enthält. Zu den Metadaten gehören:

- Meldungsnummer
- Publikationsdatum
- Heading
- Subheading
- Tenant
- Title

In einem ersten Schritt wurde ein Crawler geschrieben, um alle Metadaten zu den Meldungen herunterzuladen und in einer MySQL Datenbank in einer Tabelle zu speichern. Anschliessend wurde mit den Metadaten zu den Meldungen eine erste Analyse erstellt. In der Analyse wurde die Verteilung der Metadaten nach dem Publikationsdatum und dem Meldungstyp untersucht um relevante Meldungen für die Untersuchung der Gründungen und Löschungen zu identifizieren.

Anschliessend wurden die relevanten Meldungen über die shab.ch API heruntergeladen [https://www.shab.ch/api/v1/archive/${notice.id}/pdf?tenant=shab] und lokal gespeichert. Um die Datenmenge zu begrenzen wurden nur die Gründungsmeldungen und die Löschungsmeldungen heruntergeladen, da diese Auskunft über die Lebensdauer geben. Meldungen, deren Download nicht erfolgreich verlief, wurden in einer Datenbank erfasst, um Auskunft über die Datenqualität geben zu können und um einen weiteren Download Versuch zu ermöglichen. Der Downloader wurde auf einem Server ausgeführt. Die heruntergeladene Meldung im PDF-Format wurden mithilfe eines NPM Packages zu Text transformiert, um eine Analyse zu ermöglichen. Der Text wurde mit der Meldungsnummer in einer Datenbank erfasst und kann mit den Metadaten zu den Meldungen verbunden werden. In der Datenbank sind die Metadaten zu sämtlichen im Zeitraum veröffentlichten Meldungen vorhanden. Für alle Gründung- und Löschungsmeldungen ist ausserdem der komplette Meldungstext vorhanden.

3.3 Aufbereitung der Meldungen

3.3.1 Text Chunking

Im ersten Schritt des Preprocessing wurden aus jeder Meldung, die als unstrukturierter Text vorhanden ist, der Titel, die ID der erwähnten Firma, die Kopfzeile und die Fusszeile extrahiert um einen semi-strukturierten Datensatz zu generieren.

Abbildung 1 Kopf ohne Informationsgehalt wird entfernt

Für jede Meldung wurden Seitenheader und andere Seitendekorationen identifiziert und entfernt wie bei Abbildung 1 ersichtlich ist. Für die Identifizierung der Dekorationen wurde jeweils für jedes Layout ein Sample genommen und mit Regex die Decoratoren identifiziert. Nach dem Entfernen der Seiten-Decoratoren besteht der Text ausschliesslich aus dem Formatiertem Meldungstext.

Untersuchungen der Meldungen haben gezeigt, dass der erste Satz der Meldung jeweils die vollständige Kopfzeile mit Informationen zum Unternehmen wie Titel, Rechtsform und Gründungsort enthält. Gefolgt von einem oder mehreren Sätzen zum Zweck und anderen, je nach Rechtsform unterschiedlichen, notwendigen Informationen für die Gründung. Bei den Löschungen wird im Satz nach der Kopfzeile der Grund der Löschung angegeben. Der letzte Satz der Meldung enthält jeweils die Fusszeile.

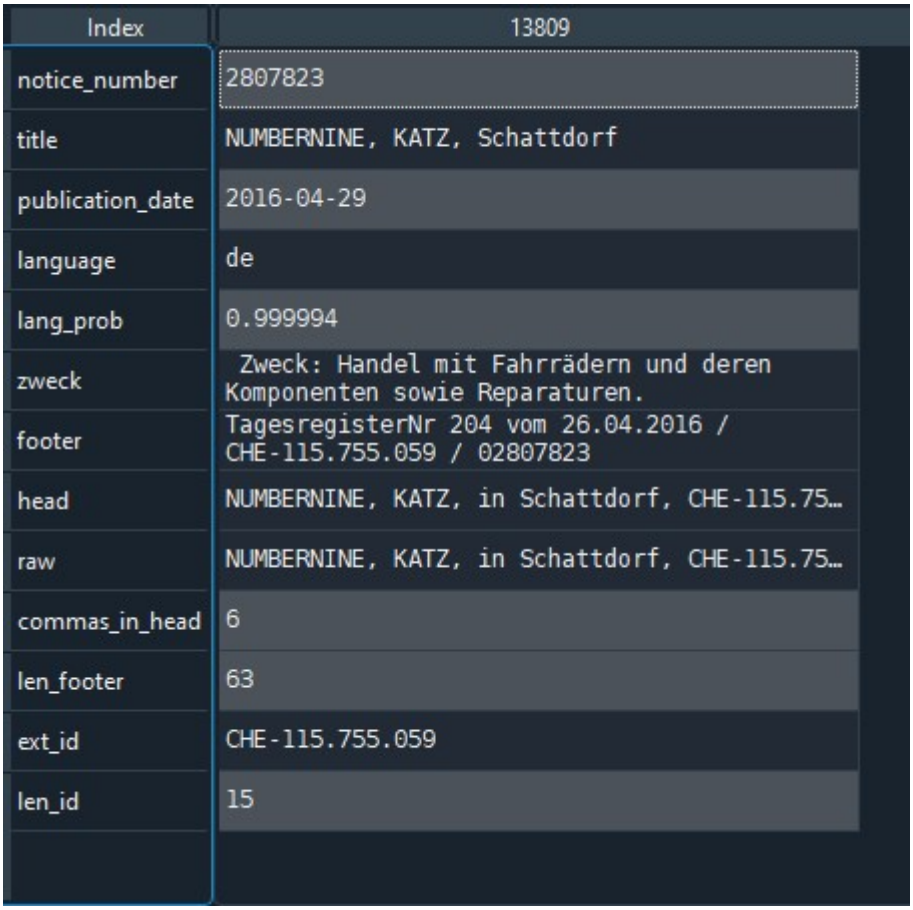
Walter Greber, Tautona Birkenzucker, in Neerach, CH-020.1. 066.329-1, Rebhaldenstrasse 12, 8173 Riedt bei Neerach, Einzelunternehmen (Neueintragung). Zweck: Handel mit Birkenzucker und natürlichen Nahrungsmitteln. Handel mit Waren aller Art. Eingetragene Personen: Greber, Walter, von Zürich, in Neerach, Inhaber, mit Einzelunterschrift; Greber-Bammatter, Ursula Maria, von Oberkirch, in Neerach, mit Einzelunterschrift.
Tagesregister-Nr. 23359 vom 23.07.2013 / CH-020.1.066.329-1 / 00999985

Abbildung 2 Aufteilung der Meldung auf Sätze

Der formatierte Meldungstext wird mit SPACY, einem Python NLP Tool in Sätze aufgeteilt. Um die Qualität der Fragmentierung in Sätze zu erhöhen, werden Formatierungen, die auch als Satztrennzeichen missverstanden werden können, identifiziert und entfernt. Zu den störenden Formatierungen gehören Silbentrennungen, Zeilenumbrüche oder neue Absätze.

Das so entstandene Array von Sätzen wird geprüft und in eine Tool-Chain übergeben, die ein Datenobjekt aus der Meldung generiert. Das Array iteriert und Kopfzeile, Fusszeile und Zweck wird identifiziert und im Datenobjekt gespeichert, ausserdem wird die in der Fusszeile enthaltene CHID oder UID extrahiert und in einem separaten Feld (ext_id) abgespeichert.

Sämtliche Datenobjekte werden transformiert und in einem Pandas Dataframe gespeichert. Abbildung 3 zeigt ein Datenobjekt, das Feature ext_id dient zur eindeutigen Identifikation des Unternehmens.



Index	13809
notice_number	2807823
title	NUMBERNINE, KATZ, Schattdorf
publication_date	2016-04-29
language	de
lang_prob	0.999994
zweck	Zweck: Handel mit Fahrrädern und deren Komponenten sowie Reparaturen.
footer	TagesregisterNr 204 vom 26.04.2016 / CHE-115.755.059 / 02807823
head	NUMBERNINE, KATZ, in Schattdorf, CHE-115.75…
raw	NUMBERNINE, KATZ, in Schattdorf, CHE-115.75…
commas_in_head	6
len_footer	63
ext_id	CHE-115.755.059
len_id	15

Abbildung 3 Beispiel eines Datenobjekts. Das Feature ext_id dient zur klaren Identifizierung des Unternehmens.

3.3.2 UID und CHID

Um die Unternehmen über den kompletten Zeitraum verfolgen zu können, muss die CHID auf die UID gemappt werden. Die zefix.ch REST API ermöglicht eine Suche nach CHID oder UID und liefert die Suchergebenisse als JSON zurück. Zefix liefert neben der CHID und der UID ebenfalls die ID des Handelsregisteramtes sowie weitere Informationen zurück. Um Informationen von zefix.ch abzufragen wurde ein Node Script erstellt, welches aus der Tabelle mit den aufbereiteten Meldungen alle CHID

und UID als Array abruft und für jede UID oder CHID einen POST Request an die API sendet. Die strukturiere Antwort im JSON wird vom Script in der Datenbank gespeichert.

3.3.3 Aufteilen der Meldungen

Zuerst wurde die Schnittmenge aller Neueintragungen und Löschungen, die im gleichen ID System bestehen, ermittelt. Für die verbliebenden Meldungen wurde die CHID auf die UID gemappt. Nach dem Mapping wurde noch einmal die Schnittmenge der verbliebenen Neueintragungen und Löschungen ermittelt. Das so entstandene Dataframe enthält alle Unternehmen die im Untersuchungszeitraum gegründet und gelöscht wurden.

Für die Untersuchung der Lebensdauer wurde eine neue Spalte im DataFrame eingefügt. Die Lebensdauer der Unternehmen wurde aus der Differenz der Publikationsdaten der Löschung und der Neueintragungen in Tagen berechnet.

Alle Neueintragungen ohne Löschung und Löschungen ohne Neueintragungen wurden in separate Dataframes gespeichert um deren Verteilung festzustellen. Es wurde eine neue Spalte eingefügt um den zeitlichen Abstand zum Start des Untersuchungszeitraumes festzustellen. Dazu wurde die Differenz zwischen dem Start des Untersuchungszeitraumes und des Publikationsdatums der Meldung in Tagen berechnet.

3.4 Clustering des Firmenzwecks

3.4.1 Erstellung TF-IDF Matrix

Um die Unternehmen anhand ihres Zwecks zu Clustern zuzuordnen wurden alle Neueintragungen in deutscher Sprache verwendet.

In einem ersten Schritt wurde aus den Datenobjekten aus Abbildung 3 das Feature **head** und das Feature **zweck** zu einem Text zusammengefügt. Anschliessend wurden alle Tags im Format &TAG; wie z.B. & für & entfernt.

Anschliessend wird der Text in einer Pipeline vorbereitet:

1. Kleinschreibung des Textes
2. Entfernen von Stoppwörtern wie: der, die, das, und
3. Entfernen der Wortsuffixe mit einem Snowball Stemmer. Aus montage wird montag

Sobald alle Texte vorbereitet sind werden sie mit einem Tfidf Vectorizer in eine TF-IDF Matrix überführt.[7] [8]

Um die Dimensionen (Anzahl Features) der TF-IDF Matrix zu reduzieren wurde der TfIdf-Vectorizer mit min_df=3 instanziiert. Der Wert min_df=3 wurde mit einer Ellbow-Heuristik festgestellt wobei min_df auf der x -Achse liegt und die Anzahl Features auf der y-Achse liegt. Die Elbow-Heuristik wir meistens für die Bestimmung der optimalen Anzahl Cluster für die KMeans Clustering Methode verwendet, ein Beispiel ist in dieser Arbeit [9] zu sehen.

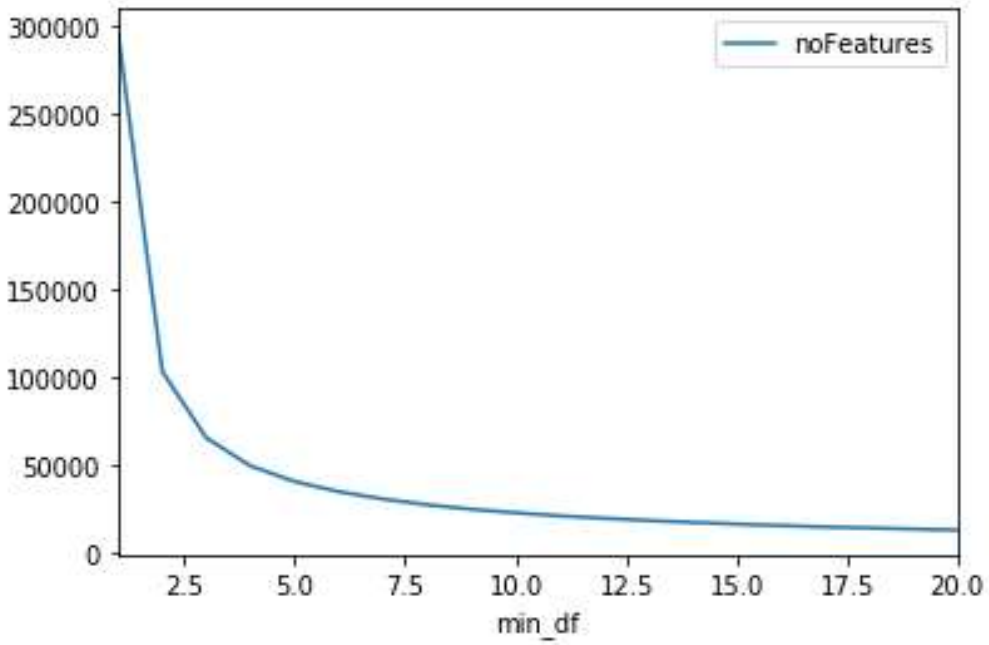


Abbildung 4Ellbow Heuristik für min_df

Die Elbow Heuristik für den minimalen Term Count zeigt das die Nummer der Features am Anfang stark fällt und die negative Steigung mit steigendem min_df abnimmt. Mit min_df = 4 sind es noch 49800 Features.

3.4.2 Reduktion der Dimensionen

Das in 3.4.1 erstellte DataFrame enthält 49800 Features und hat somit viele Dimensionen. Hochdimensionale Datensätze erschweren das Clustering[10], weil sie den Dimensionsfluch verursachen und die Effizienz der Algorithmen verringern. [8] Novembre et al. (2008) haben im ihrem Paper PCA verwendet um die Dimensionen ihres Samples massiv zu reduzieren.[11] Zusätz-

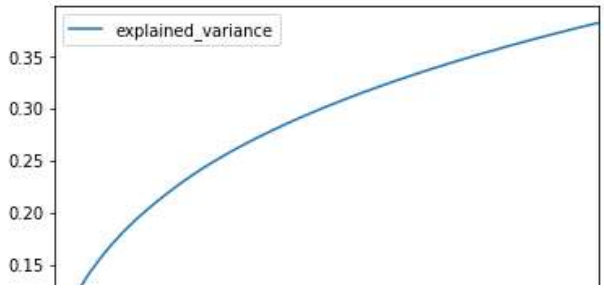


Abbildung 5 Logarithmische Zunahme der erklärten Varianz

lich wird durch die Dimensionsreduktion der Rechenaufwand für das Cluster, insbesondere die Berechnung der Silhouette Score, verringert. Die TF-IDF Matrix wird mit dem Truncated SVD Verfahren auf 1000 Features reduziert. Versuche zeigten, dass die erklärte Varianz mit der Anzahl Features logarithmisch zunimmt und der Zeitbedarf linear zunimmt.

Jedoch sind die Ressourcen der Arbeitsstation begrenzt und damit die Anzahl der Features durch den verfügbaren Arbeitsspeicher limitiert. Ausserdem erhöht die Anzahl der Features die Berechnungszeit für das Clustering. Nach der Reduzierung wird die TF-IDF normalisiert um die Berechnung für die Cluster zu vereinfachen.

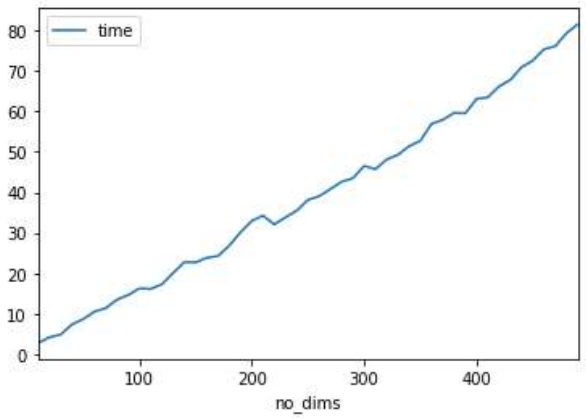


Abbildung 6 Lineare Zunahme der benötigten Rechenzeit

3.4.3 Clustering mit KMeans

Zur Bestimmung der Anzahl Cluster Center wird auf die NOGA Codes zurückgegriffen. Versuche, die Anzahl Cluster Center mit der Ellbow-Heuristik zu bestimmen, lieferten unbefriedigende Ergebnisse, siehe Kap 4.2. Als Anzahl Cluster Center wurde 21 festgelegt, was der Anzahl der NOGA Abschnitte entspricht. Um die Rechenzeit zu verkürzen wird das MiniBatch KMeans Verfahren mit der Initialisierungsmethode KMeans++ verwendet. Wie Celebi et al. in ihrem Paper [5] empfohlen haben, wird eine nicht deterministische Initialisierung mit mehreren Iterationen verwendet. Die Clustering Methode wird mit 20 Wiederholungen aufgerufen und mit der Silhouette Score bewertet. Die Initialisierung mit der höchsten Score wird für die Definition der Cluster verwendet.

Jedes Objekt im DataFrame wird einem Cluster zugeordnet. Für jedes Cluster wird eine Wordcloud erstellt, um das Cluster zu identifizieren. Um die Wordcloud zu bilden wurden aus allen Objekten im gleichen Cluster der Text verkettet und aus diesem Gesamttext mit dem Python Package Wordcloud ein Cluster gebildet. Um die Identifizierung zu vereinfachen werden Wörter, die in allen Clustern, vorhanden sind entfernt.

Die Entfernung dieser Wörter reduzierte die Menge der Worte pro Cluster um bis zu 77.8%.

Tabelle 1 Reduzierte Cluster

Cluster	# Wörter	# Wörter reduziert	Differenz	Differenz %
0	81’342	38’889	42’453	52.19%
1	151’381	65’092	86’289	57.00%
2	192’753	64’304	128’449	66.64%
3	896’692	214’037	682’655	76.13%
4	96’820	43’084	53’736	55.50%
5	409’707	125’373	284’334	69.40%
6	442’226	177’637	264’589	59.83%
7	196’392	99’069	97’323	49.56%
8	28’114	12’955	15’159	53.92%
9	210’069	84’575	125’494	59.74%
10	208’906	49’572	159’334	76.27%
11	372’471	181’645	190’826	51.23%
12	67’113	23’572	43’541	64.88%
13	380’809	84’471	296’338	77.82%
14	121’593	49’991	71’602	58.89%
15	120’538	45’625	74’913	62.15%
16	488’813	125’882	362’931	74.25%
17	268’602	110’920	157’682	58.70%
18	589’041	234’169	354’872	60.25%
19	967’116	621’603	345’513	35.73%
20	199’568	98’026	101’542	50.88%

Alle Wordclouds wurden von Hand untersucht und die Cluster identifiziert. Unklare Cluster wurden der Kategorie nicht zuordbar zugeordnet. Cluster in denen kleinere

Unternehmen wie z.B. Pronto Tankstellen, Kioske oder Coiffeursalon erwähnt werden, wurden der Kategorie Kleine Unternehmungen zugeordnet.

Abbildung 7 zeigt ein Beispiel einer Wordcloud. Häufig erwähnte Begriffe in der Wordcloud sind *software, hardware, informat* usw. Die Wordcloud zeigt einen starken Bezug zu Informationsdienstleistungen und wurde der Kategorie ID zugeordnet.

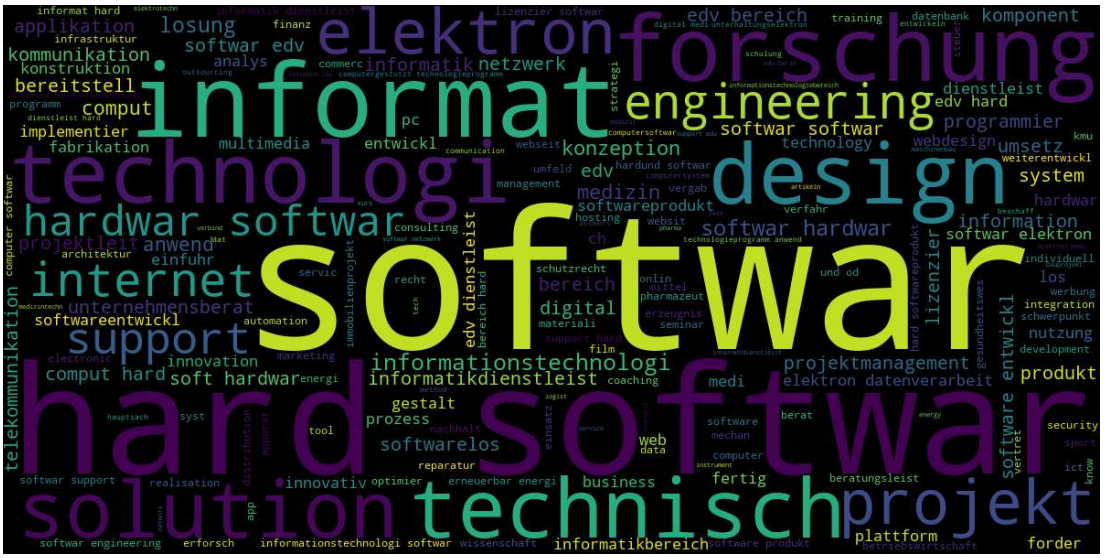


Abbildung 7 Wordcloud mit starkem Bezug zu IT

Die kategorisierten Objekte wurden anschliessend über die ID mit den bereits erstellten Gründungsmeldungen und Lebenszeiten Dataframes gemerged, um Aussagen über die Lebenszeit und die Eintrittsrate machen zu können.

3.5 Vergleich Unternehmenstext und NOGA Beschreibung

Neben dem Clustering wurden für die Untersuchung die Unternehmenstexte gegen Stichworte aus dem NOGA Verzeichnis[12] verglichen. Um die Stichworte zu generieren, wurden aus der Beschreibung der NOGA Gruppen sämtliche Worte, die ausschliesslich in den Kategorien Informatik (Informationsdienstleistungen) vorkommen, extrahiert. Die Stichworte wurden danach von Hand auf offensichtliche Fehlwörter untersucht und diese entfernt. Anschliessend wurden die Stichworte ebenfalls mit einem Snowball Stemmer gestemmt und als Keywords in einem Array gespeichert.

Für jedes Objekt wurde geprüft, ob eines oder mehrere Keywords im Text vorkommen. Falls dies der Fall war, wurde dem Objekt das Label IT-Unternehmen zugeordnet.

3.5.1 Generierung Stichwort Index aus NOGA Erläuterungen

Das Bundesamt für Statistik verwendet die NOGA Codes um die Unternehmen in statische Einheiten zu klassieren. (Allgemeine Systematik der Wirtschaftszweige NOGA)

Für die Generierung wurde aus dem Dokument «NOGA 2008 Allgemeine Systematik der Wirtschaftszweige Erläuterungen»[12] mit einem Node JS Script der reine Text extrahiert und in 2 disjunkte Teile aufgeteilt. Der Teil *Allgemein* enthält alle Erläuterungen zu den Gruppen, ausser die Erläuterungen zu den Abteilungen 62 *(Erbringung von Dienstleistungen der Informationstechnologie)* und 63 *(Informationsdienstleistungen)*. Der Teil *IT-Erklärungen* enthält ausschliesslich die Abteilungen 62 und 63.

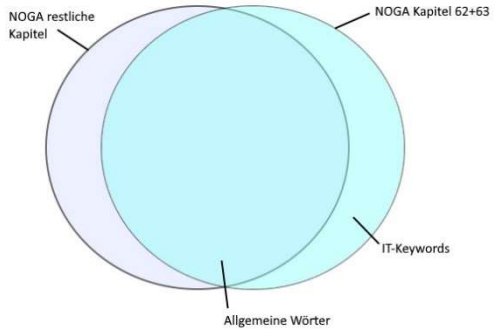


Abbildung 8 Identifizierung der IT-Keywords

Anschliessend wurden aus beiden Teilen je ein Array aus Worten gebildet. Jedes Wort, das in beiden Arrays vorhanden ist, wurde aus dem Array IT-Erklärungen entfernt. Das Array mit den IT-Erklärungen wurde anschliessend von Hand bereinigt und Worte, wie z.B. nachfolgend, die nicht direkt mit IT in Zusammenhang stehen, entfernt. Anschliessend wurden die IT-Erklärungen kleingeschrieben, von Stopwörtern bereinigt und mit dem SnowBall Stemmer gestemmt.

3.5.2 Vergleich IT-Erläuterungen und Unternehmenstexte

Dem DataFrame wurde ein neues, binäres Feature hinzugefügt, das beschreibt ob ein Stichwort der IT-Erläuterungen im Unternehmenstext erwähnt wird. Dazu wurde mit einer Lambda Funktion für jedes Unternehmen geprüft ob ein Stichwort im Unternehmenstext vorkommt, falls ja, wird dem Label der Wert 1 gegeben, falls nein, 0. Anschliessend wurden das Label auf die in 3.3 erstellten DataFrames gemappt um die Lebenszeit und die Gründungsfrequenz der IT Firmen zu beurteilen.

4 Methoden ohne befriedigende Ergebnisse

4.1 Preprocessing mit Node JS

In einem ersten Versuch wurde mit Node JS und Regex das Text Chunking durchgeführt, der Versuch wurde aufgrund der vielen Text Variationen und dem hohen Bedarf

an Rechenzeit abgebrochen. Ausschlaggebend für den Abbruch war ausserdem der Zeitbedarf, der benötigt wurde um die Satztrennzeichen zu identifizieren ohne spezielle Formatierungen wie z.B. St. Gallen zu entfernen. Nach einem Gespärch mit einer Expertin wurde das Vorgehen mit Python und Spacy verfolgt, was zu besseren Ergebnissen in kürzerer Zeit führte.

4.2 Ermittlung der IT Firmen mit Semi-Supervised Learning

Karlos et al. zeigten, dass SSL Ansätze zu guten Ergebnissen führen können[13]. Im Versuch wurde ein Sample mit 1000 Datensätzen von Hand bewertet und anschliessend mit Naive Bayes eine erste Klassifizierung durchgeführt. Nach der ersten Iteration wurden die Labels mit einer Wahrscheinlichkeit nahe an 0.5 von Hand klassifiziert. Der Versuch wurde abgebrochen da der Zeitaufwand für das Verfahren sowie die Ergebnisse nach den ersten Iterationen unbefriedigend waren. Eine Beschreibung der Naive Bayes Implementierung von SciKit-Learn zeigt ausserdem das bei Naive Bayes die Wahrscheinlichkeiten wenig Aussagekraft haben.[14]

4.3 Fehlgeschlagener Versuch KMeans Clusterermittlung

Um die optimale Anzahl der Cluster zu ermitteln wurde anhand eines Samples mit n = 1000 eine Elbow-Heuristik[9] verwendet; zusätzlich wird die Initialisierung der Cluster mit 99 Repetitionen initialisiert um die optimale Verteilung der Initialen Cluster Center zu erreichen. Die Silhouette Score und die Anzahl Cluster wurde in einem Graph dargestellt und die Anzahl Cluster wurden festgelegt.
Der Versuch schlug fehl da die Elbow-Heuristik mit jeder Iteration den Knick an einer anderen Stelle hatte und somit keine Reproduzierbaren Ergebnisse lieferte und die benötigte Rechenzeit hoch war.

5 Ergebnisse

5.1 Meldungen

Im Zeitraum von 16.01.2002 bis zum 31.08.2018 konnten mit dem in Kap 3.2 erläuterten Crawler 3'669’763 Meldungen gesammelt werden, eine manuelle Suche im Shab.ch Archiv zeigt mit den gegebenen Suchkriterien Handelsregister im Zeitraum 16.01.2002 bis zum 31.08.2018 ebenfalls 3'669'763 Meldungen. Es wurden somit alle Metadaten zu den Meldungen heruntergeladen.

Die zeitliche Verteilung der Meldung summiert auf den Tag zeigt, dass insgesamt an 4187 Tagen Meldungen publiziert wurden. Am wenigsten Meldungen wurden am 06.01.2011 mit 136 Meldungen publiziert. Am 06.04.2017 wurden mit 2563 Meldungen am meisten Meldungen publiziert. Im Durchschnitt wurden pro Tag 876.4460 Meldungen publiziert.

Die Verteilung nach Typen zeigt, dass es 3 verschiedene Typen von Meldungen mit Codes gibt. Meldungen mit dem Code hr01 stehen für Neueintragungen, hr02 steht für Mutationen bestehender Firmen und hr03 für Löschungen der Firmen. Von den insgesamt 3'669'763 Meldungen sind 2'614'203 Mutationen, 625'291 Neueintragungen und 430'269 Löschungen. Abbildung 9 zeigt visualisiert die Anteile der 3 Meldungstypen.

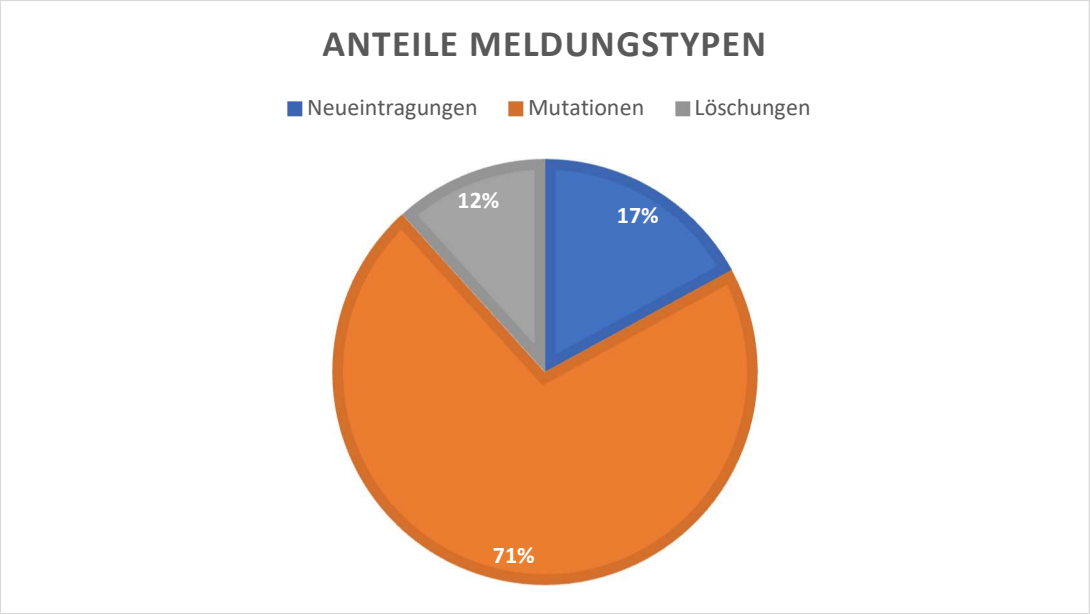


Abbildung 9 Anteile der Meldungstypen

Mit 71% sind die Mutationen am meisten vertreten gefolgt von den Neueintragungen mit 17% und Löschungen mit 12%. Der 5% höhere Anteil an Neueintragungen bedeutet, dass im Untersuchungszeitraum die effektive Anzahl von Unternehmen um 193'233 zugenommen hat.

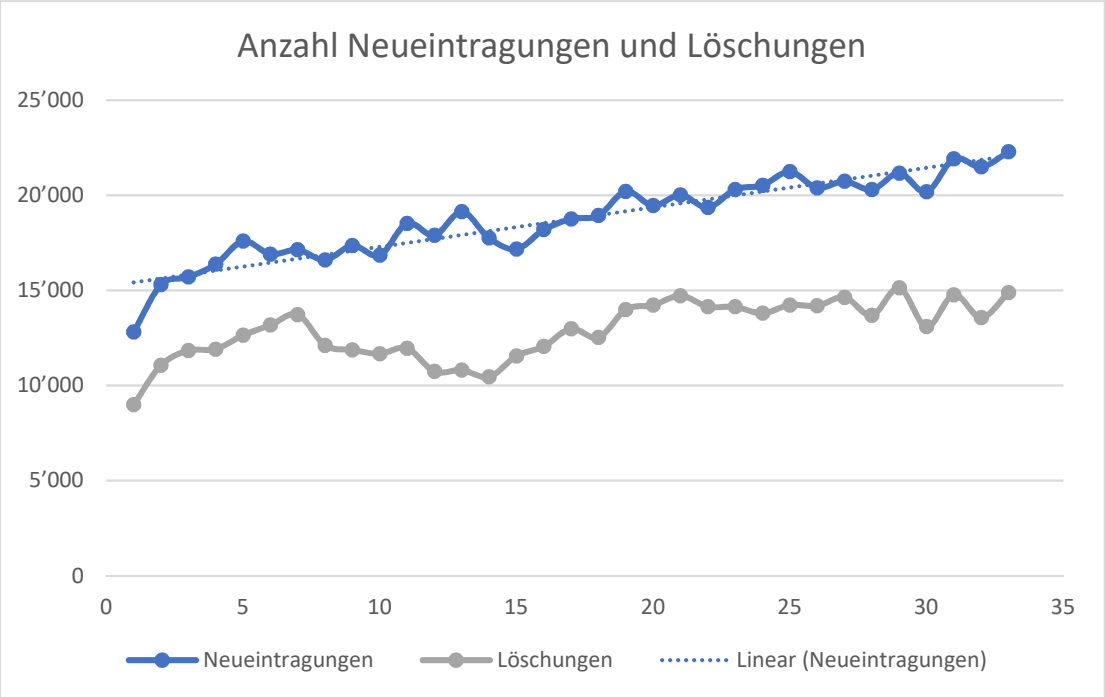


Abbildung 10Trend der Anzahl Neueintragungen und Löschungen

Abbildung 10 zeigt das die Anzahl der Neueintragungen im Trend linear zunimmt und immer in jedem Halbjahr höher als die Anzahl der Löschungen ist. Abbildung 11 zeigt die Anzahl als Balkendiagramm visualisiert.

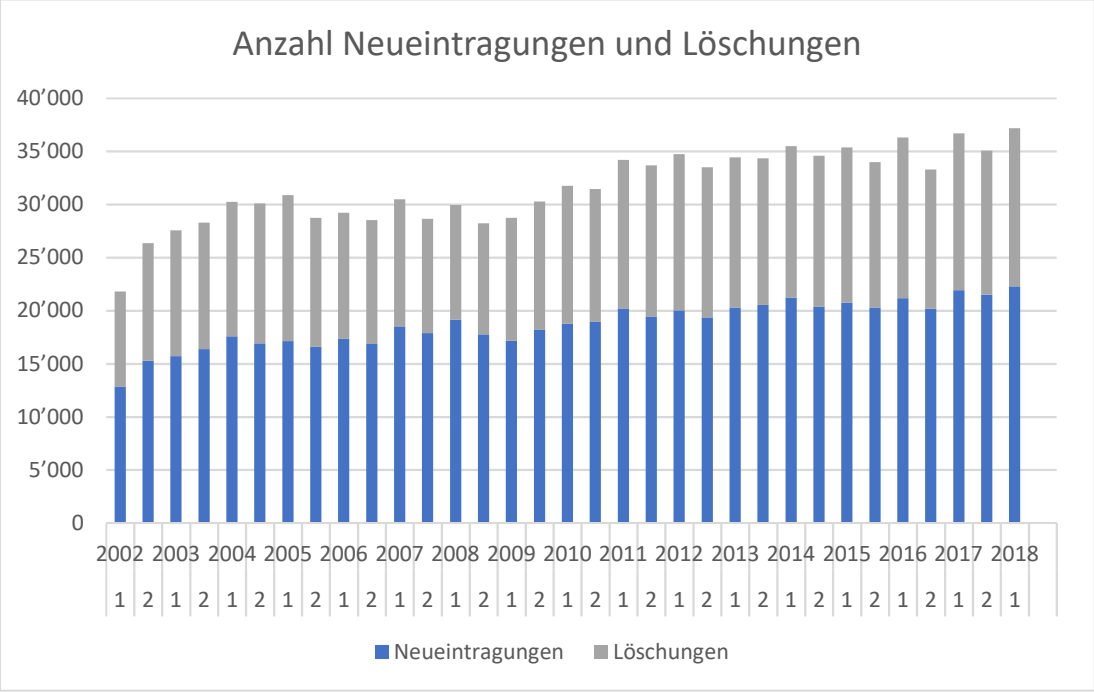


Abbildung 11 Anzahl Neueintragungen und Löschungen

Die kombinierte Verteilung der Meldungen nach Datum und Typ, aggregiert zu Halbjahren zeigt sowohl bei den Gründungen als auch bei den Löschungen einen steigenden Trend. Die Gründungen wachsen stärker als die Löschungen was dazu führt das der Zuwachs der Firmen pro Halbjahr im Trend steigt, wobei es Schwankungen gibt wie in Abbildung 12 ersichtlich ist.

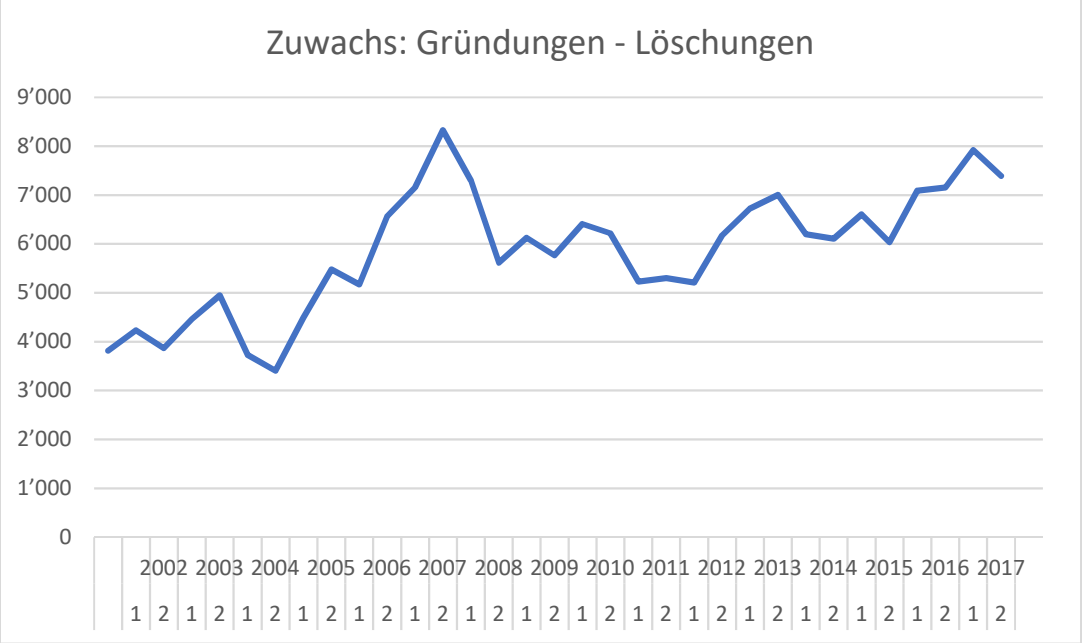


Abbildung 12 Zuwachs der Meldungen

5.2 Dynamik IT-Firmen nach NOGA Codes

Für die Untersuchung wurden nur die deutschsprachigen Neueintragungen betrachtet. Von den 625'291 Neueintragungen sind 408'467 in deutscher Sprache verfasst. 25'158 Firmen wurden mit den NOGA-Keywords als IT-Firmen klassifiziert was 6.16% entspricht.

Die Verteilung der Lebenszeiten von IT-Unternehmen und der nicht-IT Unternehmen zeigt, dass beide Kategorien einer exponentiellen Verteilung folgen.

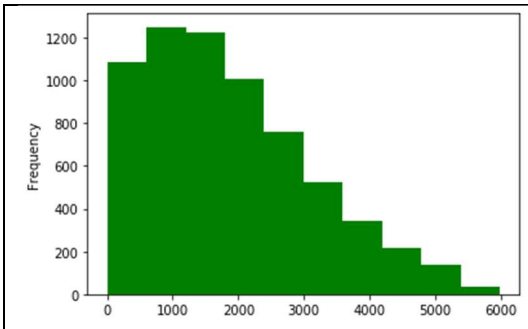


Abbildung 13 Histogramm der Lebenszeiten von IT-Firmen

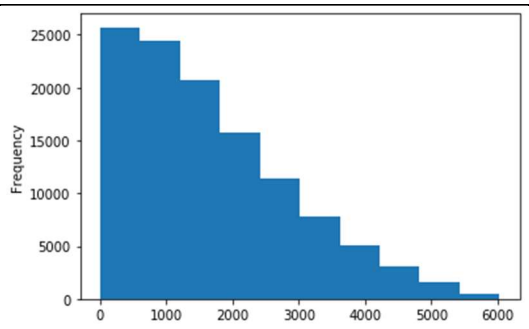


Abbildung 14 Histogramm von Lebenszeiten von nicht IT-Firmen

Allerdings zeigt das Histogramm der Lebenszeit von IT-Unternehmen den Modus bei 1000 Tagen Lebenszeit, während der Modus von nicht-IT Firmen bei 0 Tagen liegt. Eine Erklärung für diesen Effekt könnte die Honeymoon Periode sein. Beim Honeymoon Effekt sind Firmen von einem plötzlichen Austritt durch ihren Anfangsstock von Ressourcen geschützt.[15]

5.3 Clustering nach Firmenzweck

5.3.1 Identifizierung Cluster anhand der Wordclouds

Mit den 21 Wordclouds konnten neben den IT-Unternehmen weitere Bereiche identifiziert werden, mit denen die IT-Unternehmen verglichen werden können. Die Cluster wurden auf einzelne Kategorien aufgeschlüsselt.

Tabelle 2 Zuordnung Cluster zu Kategorien

Nummer	Bereich	Cluster	Anzahl Neueintragungen
0	Gemeinnützige Unternehmen	0	4’319
1	Nicht zuordbar	1, 4, 5, 6, 7, 8, 11, 15	158’580
2	IT-Bereich	2	14’379
3	Gastronomie	9	13’339
4	Reinigung	10	12’364
5	Gewerbe	12	8’434
6	Beratungen	13, 20	27’364

7	Kleine Unternehmungen, z.B. Kiosk, Coiffeur	3, 14, 18, 19	124’074
8	Handel	16	23’896
9	Transporte	17	21’718

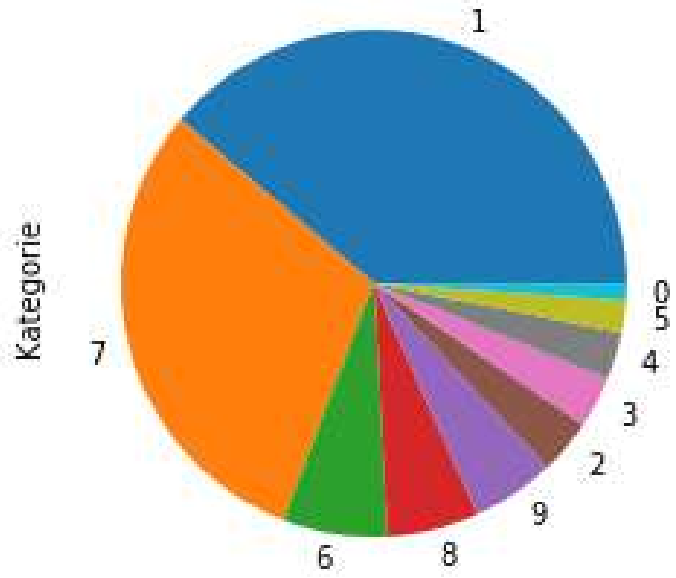


Abbildung 15 Anteil der Kategorien an allen Neueintragungen

Auf dem Kuchendiagramm wird ersichtlich, dass der Anteil, der nicht klar zuordbaren Objekten (1) am grössten ist, gefolgt von den kleineren Unternehmen (7). Bei den kleineren Kategorien führen die Beratungsunternehmen (6), gefolgt von Handel (8) und Transporte (9) mit jeweils noch über 20'000 Objekten. Die Kategorien IT-Unternehmen (2), Gastronomie (3) und Reinigung (4) haben über 10'000 Einträge und bildeten bei den Wordclouds die klarste Zuordnung. Unter 10'000 Einträge fallen die Kategorie Gewerbe (5) und Gemeinnützige Unternehmen (0). Die Kategorie Gemeinnützige Unternehmen bildet mit 4'319 Objekten die kleinste Kategorie.

5.3.2 Verteilung der Lebenszeiten IT, Gastronomie und Reinigung

Die Kategorien IT, Gastronomie und Reinigung haben mit 14'379, 13'339 und 12'364 Neueintragungen ähnlichen Anteile an den insgesamt 408'467 Neueintragungen und werden aufgrund der ähnlichen Anteile miteinander verglichen.

Tabelle 3 Vergleich der 3 Kategorien

Kategorie	Anzahl Neueintritte	Gelöscht	Existierend	Überlebensrate
IT-Unternehmen	14’379	6’521	7’858	0.55
Gastronomie	13’339	3’615	9’724	0.73

Reinigung	12’364	2’393	9’971	0.81

Der Vergleich der Überlebensrate zeigt, dass IT-Unternehmen eine tiefere Überlebensrate haben.

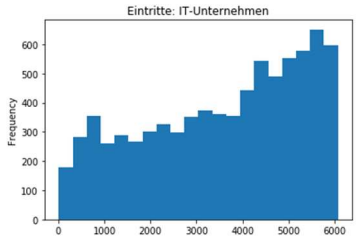


Abbildung 16 Eintritte der noch existierenden IT-Unternehmen

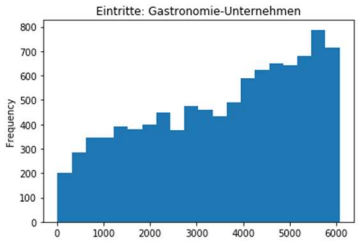


Abbildung 17 Eintritte der noch existieren-den Gastronomie-Unternehmen

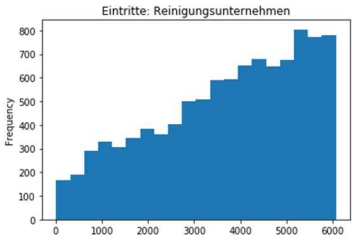


Abbildung 18 Eintritte der noch existierenden Reinigungs-Unternehmen

Die Verteilungen in den Abbildungen 13, 14, 15 zeigen die Eintritte der noch existierenden Unternehmen. Die Eintritte sind steigend, weil der Untersuchungszeitraum begrenzt ist und Neueintritte am Ende des Untersuchungszeitraum weniger Austrittsmöglichkeiten hatten als Neueintritte am Anfang des Untersuchungszeitraumes. Die Verteilung der IT-Eintritte zeigt einen Knick bei 4000 Tagen nach Untersuchungsstar am 16.01.2002, ab 2013 steigt die Anzahl der überlebenden IT-Eintritte stärker. Die Eintritte der Gastronomie und Reinigungsunternehmen zeigen weniger ausgeprägte Knicke.

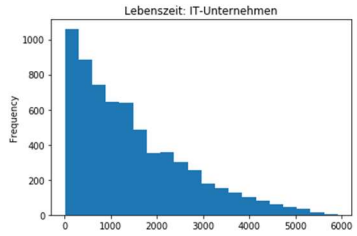


Abbildung 19 Verteilung der Lebenszeit von IT-Unternehmen

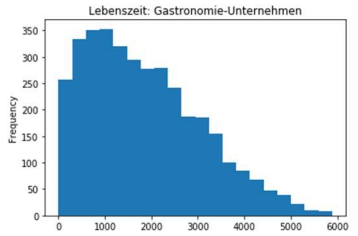


Abbildung 20 Verteilung der Lebenszeit von Gastronomie-Unternehmen

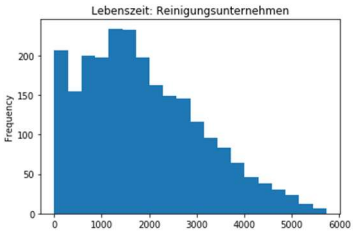


Abbildung 21 Verteilung der Lebenszeit von Reinigungs-Unternehmen

Die Abbildungen 16, 17, 18 zeigen die Verteilung der Lebenszeiten der verschiedenen Kategorien. Die Verteilung der Lebenszeit von IT-Unternehmen hat den Modus bei 0 Tagen und fällt mit der Dauer. Anders als die in Kap. 5.2 ermittelten Unternehmen hat die Verteilung keine Honeymoon Periode. Die Verteilung der Lebenszeiten der Gastronomie- und Reinigungsunternehmen haben den Modus bei 1000 Tagen oder später.

6 Diskussion

6.1 Zusammenfassung

In der dieser Bachelor Arbeit wurden Datenquellen für Handelsregister gesucht und danach aus 1 - Stern Open Data[16], den heruntergeladenen PDF von SHAB.ch, und 4 - Stern Open Data[16], den gecrawlten JSON aus zefix.ch, Informationen generiert, um Unternehmen zu modellieren, zu kategorisieren und deren Firmendynamik zu Untersuchen. Speziell wurde der Fokus auf IT, Gastronomie und Reinigungsunternehmen gelegt da diese ähnlichen Anteile an den Neueintragungen hatten.

Die Analyse der Meldungen zeigt, dass es mehr Gründungen als Löschungen gibt und die Anzahl der Firmen in der Schweiz wächst. Die Verteilung der Lebensdauer der Neueingetretenen und Ausgetretenen folgt der von Coad vorgeschlagenen Exponentialverteilung. [2]

Der Vergleich der 3 Kategorien zeigt, dass sich nicht alle Kategorien gleich verhalten und der Sektor IT gegenüber von Gastronomie und Reinigungsunternehmen eine höhere Mortalität aufweist.

Anders als die in Kap 5.2 ermittelten IT-Neueintritte zeigen die durch KMEANS Clustering ermittelten Eintritte aus 5.3 keine Honeymoon-Periode.

6.2 Diskussion Ergebnisse

Die Ergebnisse zeigen, dass die von COAD angenommen Exponentialverteilung für die Lebensdauer, der Unternehmen die im untersuchten Zeitraum gegründet und gelöscht wurden, passend ist. Jedoch sind nur 1/5 der gegründeten Unternehmen auch wieder gelöscht worden. Ausserdem stimmt die von Coad getroffene Annahme, dass die Eintrittsrate konstant ist [2], nicht mit dem Trend der Eintrittsrate aus Abbildung 10 überein.

Die unterschiedlichen Methoden um die IT-Unternehmen zu ermitteln zeigten, dass es Unterschiede in der Verteilung der Lebenszeiten gibt. Diese könnten daraus entstehen, dass Meldungstexte in denen Keywords vorhanden noch andere Wörter vorkamen, die mehr den Kategorien 1 *nicht zuordbar* oder 6 *Beratungen* entsprachen und die Unternehmen diesen Clustern zugeordnet wurden.

6.3 Limiten

Die Datenanalysen wurden auf Basis der Daten von den als pdf archivierten EHRA Meldungen aus dem Zeitraum vom 01.01.2002 bis zum 31.08.2018 gemacht. Die Meldungen kommen aus 3 Sprachregionen. Um die Analyse zu vereinfachen wurden in dieser Arbeit nur die Meldungen in deutscher Sprache untersucht was 65.3 % der Meldungen entspricht. Mit dem Dazunehmen der französischen und italienischen Meldungen könnte der Informationsgehalt der Untersuchung erhöht und vor allem Unterschiede in den Sprachregionen festgestellt werden.

Das parsen der .pdf-Meldungen führte in einer Teilmenge zu Fehlformatierungen im Text, das Python Tool spacy wurde eingesetzt um Fehlformatierungen zu kompensieren und Sätze zu erkennen um daraus die Gründungsmeldungen zu modellieren. Je nach Fehlformatierung sind die Resultate der spacy-Untersuchung von unterschiedlicher Qualität, mit angepassten Funktionen konnte ein grosser Teil der Meldungen analysierbar gemacht werden. Jedoch ist nicht sicher, ob die Daten für alle 600'000 Gründungen und 400'000 Löschungen von guter Qualität sind.

Für Untersuchungen im Zeitraum ab 01.09.2018 sind die Meldungen im .xml Format verfügbar bei denen der Unternehmenszweck mit einem Tag identifizierbar ist. Die Datenqualität wäre dann messbar höher.

Bei der Umstellung des Shab Web Auftritts wurden alle Meldungen vor dem 31.08.2019 als .pdf archiviert und somit der Zugriff auf die darin enthaltenen Informationen erschwert. Es ist zu prüfen ob eine tägliche Abfrage auf die shab API nötig ist um die Meldungsdaten als xml erhalten zu können.

Der Hochdimensionale Datensatz erschwerte das Clustering und wurde mit dem TruncatedSVD Verfahren reduziert, was die erklärte Varianz in den Daten begrenzte und zu einem Informationsverlust führt. Um die Dimensionen mit weniger Informationsverlust zu reduzieren, könnte man versuchen die Synonyme mit einem Wörterbuch zu reduzieren.

6.4 Ausblick

In weiteren Untersuchungen sollte neben dem Eintrag ins Handelsregister ebenfalls die Umsätze der Unternehmen oder die Anzahl der Beschäftigten berücksichtigt werden um zu prüfen ob es inaktive Unternehmen gibt, die keine Wirtschaftliche Tätigkeit verfolgen, aber noch nicht aus dem Handelsregister gelöscht wurden. Weiter könnte

das Clustering mit der Einführung von Wörterbüchern für Synonyme verbessert werden. Zusätzlich könnten die Meldungen in der französischen und italienischen Sprache der Untersuchung hinzugefügt werden um Unterschiede zwischen den Sprachregionen festzustellen und ein Gesamtbild über die Schweizer Firmendynamik zu erstellen. Durch die neue Webseite von Shab.ch würden die Neueintragungen und Löschungen täglich verfügbar womit man ein Monitoring umsetzen könnte.

Abbildungsverzeichnis

Tabellenverzeichnis

Literaturverzeichnis

[1] E. Santarelli und M. Vivarelli, „Entrepreneurship and the Process of Firms Entry, Survival and Growth“, *Ind. Corp. Change*, Bd. 16, S. 455–488, Mai 2007, doi: 10.1093/icc/dtm010.

[2] A. Coad, „Investigating the Exponential Age Distribution of Firms“, *Econ. Open-Access Open-Assess. E-J.*, Bd. 4, Nr. 2010–17, S. 1, 2010, doi: 10.5018/economics-ejournal.ja.2010-17.

[3] M. I. G. Daepp, M. J. Hamilton, G. B. West, und L. M. A. Bettencourt, „The mortality of companies“, *J. R. Soc. Interface*, Bd. 12, Nr. 106, S. 20150120, Mai 2015, doi: 10.1098/rsif.2015.0120.

[4] A. Coad und C. Guenther, „Age, diversification and survival in the German machine tool industry, 1953-2002“, S. 30.

[5] G. Barba Navaretti, D. Castellani, und F. Pieri, „Age and firm growth: evidence from three European countries“, *Small Bus. Econ.*, Bd. 43, Nr. 4, S. 823–837, Dez. 2014, doi: 10.1007/s11187-014-9564-6.

[6] A. Coad, „Firm age: a survey“, *J. Evol. Econ.*, Bd. 28, Nr. 1, S. 13–43, Jan. 2018, doi: 10.1007/s00191-016-0486-0.

[7] „sklearn.feature_extraction.text.TfidfVectorizer — scikit-learn 0.22.1 documentation“. [Online]. Verfügbar unter: https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html?highlight=tfidf#sklearn.feature_extraction.text.TfidfVectorizer. [Zugegriffen: 26-Jan-2020].

[8] P. Bafna, D. Pramod, und A. Vaidya, „Document clustering: TF-IDF approach“, 2016, S. 61–66, doi: 10.1109/ICEEOT.2016.7754750.

[9] L. Recalde und R. Baeza-Yates, *What kind of content are you prone to tweet? Multi-topic Preference Model for Tweeters*. 2018.

[10] „Clustering high-dimensional data“, *Wikipedia*. 19-Dez-2019.

[11] J. Novembre *u. a.*, „Genes mirror geography within Europe“, *Nature*, Bd. 456, Nr. 7218, S. 98–101, Nov. 2008, doi: 10.1038/nature07331.

[12] B. für Statistik, „Allgemeine Systematik der Wirtschaftszweige - NOGA 2008 - Erläuterungen | Publikation“, *Bundesamt für Statistik*, 15-Dez-2008. [Online]. Verfügbar unter: /content/bfs/de/home/statistiken/industrie-dienstleistungen/nomenklaturen/noga/publikationen-noga-2008.assetdetail.344101.html. [Zugegriffen: 26-Jan-2020].

[13] S. Karlos, N. Fazakis, A.-P. Panagopoulou, S. Kotsiantis, und K. Sgarbas, „Locally application of naive Bayes for self-training“, *Evol. Syst.*, Bd. 8, Nr. 1, S. 3–18, März 2017, doi: 10.1007/s12530-016-9159-3.

[14] „1.9. Naive Bayes — scikit-learn 0.22 documentation“. [Online]. Verfügbar unter: https://scikit-learn.org/stable/modules/naive_bayes.html#complement-naive-bayes. [Zugegriffen: 20-Dez-2019].

[15] A. Coad, A. Segarra, und M. Teruel, „Like milk or wine: Does firm performance improve with age?“, *Struct. Change Econ. Dyn.*, Bd. 24, S. 173–189, März 2013, doi: 10.1016/j.strueco.2012.07.002.

[16] „5-Sterne Offene Daten“. [Online]. Verfügbar unter: http://5stardata.info/de/. [Zugegriffen: 26-Jan-2020].

Selbständigkeitserklärung

„Ich erkläre hiermit, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetzes vom 5. September 1996 über die Universität zum Entzug des aufgrund dieser Arbeit verliehenen Titels berechtigt ist.“

Bern, 31.01.2020 Dominic Schweizer

Veröffentlichung der Arbeit

(nur für Master- / Lizentiats- /Bachelorarbeit)

I.d.R. werden schriftliche Arbeiten in der Bibliothek des Instituts für Wirtschaftsinformatik öffentlich zugänglich gemacht.

- ☐ Hiermit erlaube ich, meine Arbeit in der Bibliothek des Instituts für Wirtschaftsinformatik zu veröffentlichen.
- ☐ Ich möchte auf eine Veröffentlichung meiner Arbeit verzichten.

Falls eine Vertraulichkeitserklärung unterschrieben wurde, ist es Sache des Studierenden, das Einverständnis des Praxispartners einzuholen. Es muss der Arbeit eine schriftliche Bestätigung des Praxispartners beigelegt werden.

Die Benotung der Arbeit erfolgt unabhängig davon, ob die Arbeit veröffentlicht werden darf oder nicht.

Bern, 31.01.2020 Dominic Schweizer