

Nutzerorientierte Entwicklung und Anwendung eines Monitoring Bot Prototyps: Chancen und Herausforderungen

Masterarbeit

zur Erlangung des akademischen Grades „Master of Science (M. Sc.)“ im
Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik,
Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät der
Leibniz Universität Hannover

vorgelegt von

Name: Suprun

Vorname: Kirill



Prüfer: Prof. Dr. Michael H. Breitner

Hannover, den 29.09.2022

Inhaltsverzeichnis

Kurzfassung	I
Abbildungsverzeichnis	IV
Tabellenverzeichnis	VI
Abkürzungsverzeichnis	VII
1 Einleitung	1
2 Theoretische Grundlagen	3
2.1 Digitale Wirtschaft	3
2.2 Data Mining	5
2.3 Web Scraping	7
2.4 Mensch-Maschine-Interaktion in der digitalen Wirtschaft.....	11
2.4.1 PACT-Framework.....	11
2.4.2 Crowdsourcing.....	15
2.4.3 Design Principles.....	17
3 Methodik	20
3.1 Literatursuche und Analyse.....	20
3.2 Prototyping.....	23
3.3 Monitoring durch Web Scraping	24
3.4 SWOT-Analyse	27
4 Theoretische Implementierung	29
4.1 Anwendungsfall 1: Availability-Monitoring.....	29
4.2 Anwendungsfall 2: Price-Monitoring.....	30
4.3 Benutzerschnittstelle	33
4.3.1 Eigenschaften des Servers.....	35
4.3.2 Serverstruktur	37
5 Nutzerorientierte Entwicklung des Prototyps	44
5.1 Benutzerschnittstelle	44
5.2 Availability-Monitoring	48
5.3 Price-Monitoring.....	55
6 Chancen und Herausforderungen	64
6.1 Web Scraping	64
6.2 Crowdsourcing	66
6.3 Availability-Monitoring	69

6.4 Price-Monitoring.....	71
7 Diskussion und Empfehlung	75
7.1 Diskussion der Ergebnisse.....	75
7.2 Beitrag für Forschung.....	79
7.3 Beitrag für Praxis.....	80
8 Limitationen und Ausblick.....	81
9 Fazit	83
Literaturverzeichnis.....	VIII
Anhang	XVI
A – Scraping Bibliotheken	XVI
B – Discord-Webhook	XVIII
C – Amazon Monitor	XX
D – Kaufland Monitor	XXI
E – NVIDIA API.....	XXII
F – NVIDIA Monitor.....	XXIII
G – Price Monitor	XXIV

1 Einleitung

Mit der Einführung neuer Technologien ist die Zahl der Internetnutzer und die von ihnen erzeugte Datenmenge enorm gestiegen (vgl. Khder 2021, S. 158), sodass sich das Internet als die größte Datenquelle des 21. Jahrhunderts bewährt (vgl. Bhardwaj et al. 2021, S. 261). Daten werden im digitalen Zeitalter für die unterschiedlichsten Zwecke erhoben und bilden einen wesentlichen Bestandteil in diversen Segmenten der akademischen und wissenschaftlichen Forschung (vgl. Sirisuriya 2015, S. 135). Auch bei Unternehmen verschiedener Arten und Behörden ist die Verarbeitung und Analyse generierter Daten aus dem Internet von großer Relevanz, da die Daten bei der Bestimmung von Verbraucherinteressen, Ausrichtung gezielter Werbung und Identifizierung von Bedrohungen unterstützen (vgl. Bhardwaj et al. 2021, S. 261). Angesichts des Umfangs, der Vielfalt und der Geschwindigkeit der im Internet verfügbaren Daten, ist deren Organisation manuell kaum möglich, weshalb Forscher auf Technologien zurückgreifen, um Aspekte der Datenerfassung zu automatisieren (vgl. Krotov et al. 2020, S. 539). Eine der populärsten Methoden ist das Web Scraping, welches als ein Technologie-Tool für die automatische Extraktion und Organisation von Daten aus dem Web zum Zweck der weiteren Analyse dieser Daten definiert wird (vgl. Krotov und Tennyson 2018, S. 170). Das Web Scraping ermöglicht die Extraktion bestimmter Elemente einer Webseite und die Bereitstellung dieser erhobenen Daten für weitere Anwendungen, wie das Monitoring von Preisänderungen oder die Analyse von Produktbewertungen (vgl. Uzun 2020, S. 61726). Ein wesentlicher Faktor dieses automatisierten Prozesses ist die Steigerung der Zeiteffizienz. Automatisierungen komplexer Abläufe bilden im Kontext der Digitalisierung neue Geschäftsmodelle und ermöglichen in diesem Zusammenhang Skalierungspotenziale (vgl. Malfent und Missbach 2019, S. 311) sowie neue wirtschaftliche und gesellschaftliche Formen der Arbeitsorganisation, wie die des kollektiv intelligenten Crowdsourcings (vgl. Hoßfeld et al. 2012, S. 204), welches gezielt in einer kollaborativen und wettbewerbsorientierten Organisation eine interaktive Leistung erbringt (vgl. Bertschek et al. 2015, S. 6). Die Integration von externen Personen und die Entstehung neuer Netzwerke fördert die unternehmerische Produktentwicklung und -innovation.

Eine Verknüpfung des Crowdsourcings mit einer für die Mitglieder verfügbaren Web-Scraping-Technologie als digitale Dienstleistung schafft ein breites Spektrum an unerforschter und moderner Anwendbarkeit. Obwohl viele Einzelaspekte des Web Scrapings bereits untersucht wurden, umfasst die wissenschaftliche Literatur meist lediglich spezifische technische Lösungen mit Erörterungen, wie das Web Scraping in einem bestimmten Kontext von Forschern genutzt werden kann. Trotz des Phänomens des sich fortsetzenden Einsatzes neuer Informationstechnologien (vgl. Bertschek et al. 2015, S. 7) und auftretender wertschöpfender Vorteile (vgl. GIZ 2021, S. 2), gibt es bisher wenige qualitative Forschungen, die weitere Möglichkeiten des Web Scrapings und potenzielle Bedrohungen in dessen Anwendung untersucht. In diesem Sinne besteht die Motivation dieser Thesis darin, die Chancen und Herausforderung bei der nutzerorientierten Entwicklung eines Web Scraping basierenden Monitoring Bots

anhand zwei unterschiedlicher Anwendungsfälle zu identifizieren und mit diesen einen hilfreichen Beitrag an die weiterführende Forschung und Praxis zu erarbeiten. Vor diesem Hintergrund lassen sich die folgenden zwei Forschungsfragen ableiten:

- (1) *Welche Chancen und Herausforderungen entstehen bei der Implementierung und Verwendung eines Monitoring Bots?*
- (2) *Welche Stärken und Schwächen werden bei unterschiedlichen Anwendungsfällen durch die Entwicklung des Monitoring Bots festgestellt?*

Die erste Forschungsfrage zielt darauf ab, wesentliche Erkenntnisse zum allgemeinen Web Scraping zu sammeln, um mit diesen die technische Umsetzung des angestrebten Monitor Bots zu realisieren und die damit verbundenen Chancen und Herausforderungen bei der Implementierung zu analysieren. Mit dem gewonnenen Wissen lässt sich das Prinzip des Web Scrapings auf weitere Anwendungen übertragen, woraus sich durch die zweite Forschungsfrage die Stärken und Schwächen in Bezug auf die Anwendungsfälle untersuchen lassen.

Um die beiden Forschungsfragen systematisch zu untersuchen und fundiert zu beantworten, ist die vorliegende Arbeit in neun Hauptkapitel gegliedert. An die Einleitung anknüpfend werden zunächst die theoretischen Grundlagen behandelt. Hierbei werden Aspekte der digitalen Wirtschaft sowie die Funktionsweisen des Data Minings und Web Scrapings durchleuchtet. Zudem werden relevante Gesichtspunkte aus der Mensch-Maschinen-Interaktion dargelegt. Im dritten Kapitel wird die Methodik zur Aufarbeitung der Ergebnisse dargelegt. Hierunter erfolgt eine systematische Literaturanalyse gemäß den Ansätzen nach Vom Brocke et al. (2015) sowie Webster und Watson (2020). Neben der Erläuterung der verwendeten Techniken des Web Scrapings erfolgt zudem eine kurze Einführung in die SWOT-Analyse, mit der die möglichen Chancen und Herausforderungen bei der Verwendung von Monitor Bots identifiziert werden sollen. In dem darauffolgenden vierten Kapitel wird die theoretische Implementierung durchgeführt und die beiden unterschiedlichen Anwendungsfälle des Availability- sowie Price Monitorings umfassend präsentiert. Anschließend werden die auf die Anwendungsfälle bezogenen Monitor Bots im fünften Kapitel unter Aspekten relevanter Gestaltungsprinzipien zur Sicherstellung der Nutzerorientierung entwickelt. Implementiert wird dies auf einer zugänglichen Benutzeroberfläche, die eine potenzielle Crowdsourcing-Plattform darstellt. Im sechsten Kapitel wird die SWOT-Analyse durchgeführt, dessen Ergebnisse anschließend im siebten Kapitel diskutiert und impliziert werden. Die Limitationen der eigenen Forschung sowie ein Ausblick auf mögliche Weiterentwicklungen werden im vorletzten Kapitel ausgeführt. Abschließend erfolgt eine Zusammenfassung der Ergebnisse.

9 Fazit

Im Rahmen der vorliegenden Arbeit werden verschiedene Monitor Bot Prototypen unter menschenzentrierten Aspekten entwickelt, um die Chancen und Herausforderung bei der Implementierung und Verwendung zu identifizieren. Anhand einer Literaturrecherche wird festgestellt, dass die vielfältige Einsatzmöglichkeit von Web Scraping in der wissenschaftlichen Literatur kaum thematisiert wird. Infolgedessen werden in dieser Studie zwei praxisnahe Anwendungsfälle untersucht, die durch eine automatisierte Datenerhebung in Form eines Availability- und Price-Monitorings diverse Potenziale sowohl für die Forschung als auch die digitale Wirtschaft bieten. Das Monitoring basiert auf Prinzipien des Web Scrapings und wird durch die drei Phasen der Webseiten-Analyse, -Crawling und Datenorganisation realisiert. Die drei Entwicklungsphasen erfordern aufgrund möglicher Modifizierung einer Webseite eine fortlaufende menschliche Überwachung, um auftretende Änderungen stets auszugleichen und die Datenextraktion fehlerfrei zu gewährleisten. Die Programmierung erfolgt mit der Programmiersprache Python und zielt darauf ab, ausgewählte Daten in vordefinierten Zeitabständen iterationsweise zu sammeln und sie gemäß den Anforderungen des Anwendungsziels weiter zu analysieren.

Auf Grundlage aktuell herrschender Trends sowie coronabedingter Lieferengpässe erfährt das Availability-Monitoring seine Relevanz und ermöglicht Anwendern, eine Mitteilung über die Verfügbarkeit eines auf einem Online-Shop angebotenen Produktes in Echtzeit zu empfangen. So kann die Chance erhöht werden, ein gewünschtes Produkt zu erwerben, welches streng limitiert oder stark nachgefragt und demzufolge üblicherweise ausverkauft ist. Eine manuelle Suche bei verschiedenen Online-Shops führt häufig zu keiner erfolgreichen Kaufmöglichkeit. Das Web Scraping schafft hierbei das Potenzial, eine Vielzahl verschiedener Webseiten zur selben Zeit zu überwachen und eine Verfügbarkeit durch eine Änderung auf der Webseite festzustellen. Das Ziel des Price-Monitorings hingegen liegt in der Speicherung und Gegenüberstellung von Preisdaten, die in mehreren unterschiedlichen Webseiten enthalten sind. Die manuelle Erfassung dieser Daten stellt einen erheblichen Zeitaufwand dar. Die Automatisierung dieses Prozesses führt neben der wirkungsvollen Zeiteffizienz zu weiteren Analysemöglichkeiten, indem die Daten strukturiert in einem gewünschten Format übersichtlich dokumentiert und in aussagekräftige Visualisierungen festgehalten werden können. Darüber hinaus könnte das Modell des Monitorings in ein Decision Support System integriert werden, welches den Benutzer bei der Entscheidungsfindung unterstützt und Handlungsalternativen vorschlägt. Implementiert werden die Bots auf einem zugänglichen Server, der eine gemeinschaftliche Crowdsourcing-Plattform darstellt und wesentliche Designprinzipien der Mensch-Maschine-Interaktion innehat. Hierin besteht die Chance, an die Anwendungsfälle anzuknüpfen und wertschöpfende Vorteile durch Diversität und verteiltem Wissen zu generieren. Die in diesem Zusammenhang analysierten Chancen und Herausforderungen sowie Stärken und Schwächen werden im Diskussionsteil reflektiert und für die Forschung und Praxis impliziert. Darauf folgend werden die

Limitationen dieser Ergebnisse, wie beispielsweise die ständige Weiterentwicklung von Webseiten oder vom Betreiber eingeführte Barrieren als Maßnahme für automatisierte Abfragen aufgezeigt und ein Ausblick auf die weitere Forschung gegeben, wie die Identifizierung bedeutsamer Aspekte für eine erfolgreiche Begegnung mit einem Monitor Bot oder die Weiterentwicklung des Web Scrapings unter Einsatz von künstlicher Intelligenz.

Die tabellarisierten SWOT-Analysen stellen basierend auf dem aktuellen Stand der Forschung und Technologie die Forschungsergebnisse dieser Arbeit dar. Sie tragen dazu bei, Schwächen und Risiken bereits vor der Entwicklungsphase eines Monitor Bots zu erkennen, diese zu reduzieren und Strategien zur Chancenoptimierung sowohl aus praktischer als auch forschungsbezogener Sicht abzuleiten.