

Practical Data Science.

WS 25/26



What to expect.

1. About snapAddy.

2. AI and Machine Learning at snapAddy.

3. Don't Believe the Hype.

4. Capstone Project.



1

About snapAddy.



Hey, we're snapAddy!



When it comes to contact data, people think of snapAddy.

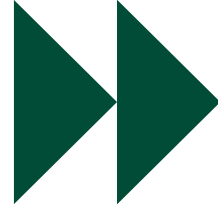
We are one of the leading software providers for automated contact capture, eliminating time-consuming manual CRM processes in marketing and sales.

Since our founding in 2015, the startup has grown into an international team with over 100 employees — and the journey continues.

More than 4,000 companies, including several DAX-listed corporations, already rely on our solutions. And new ones are joining every day.



Product Overview.



BusinessCards

Your digital business card –
always at hand.

DataQuality

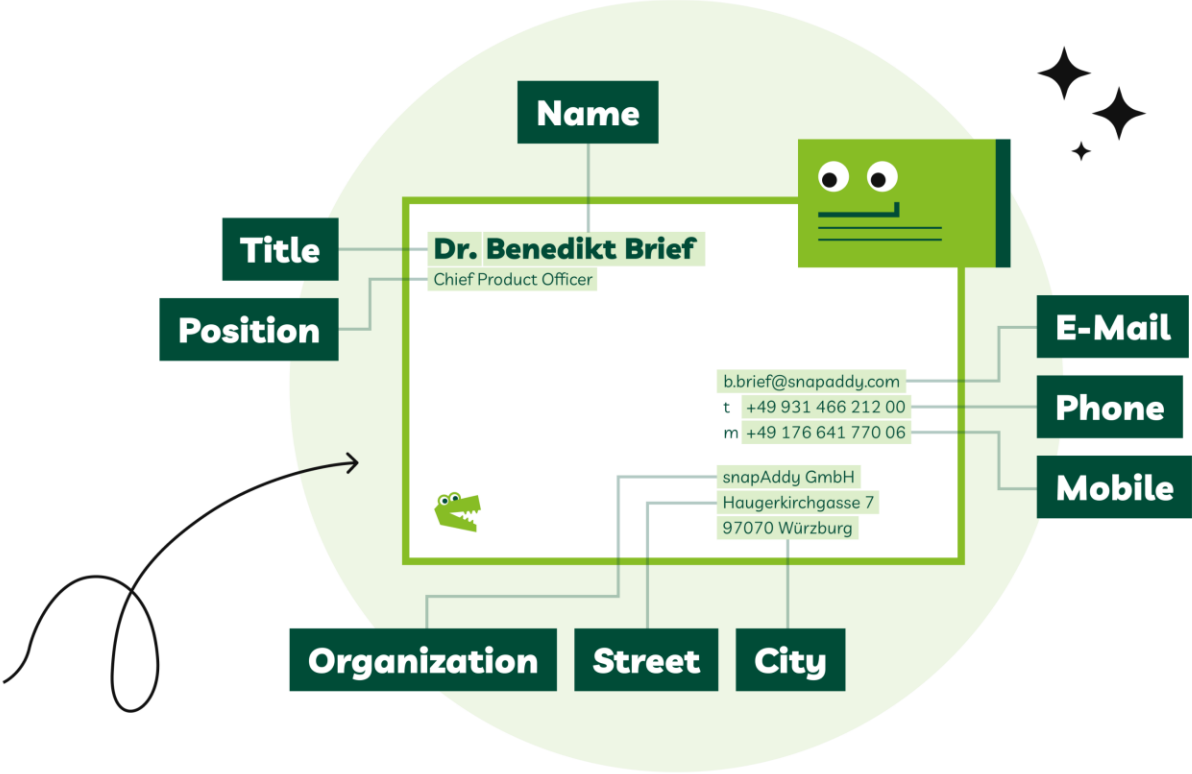
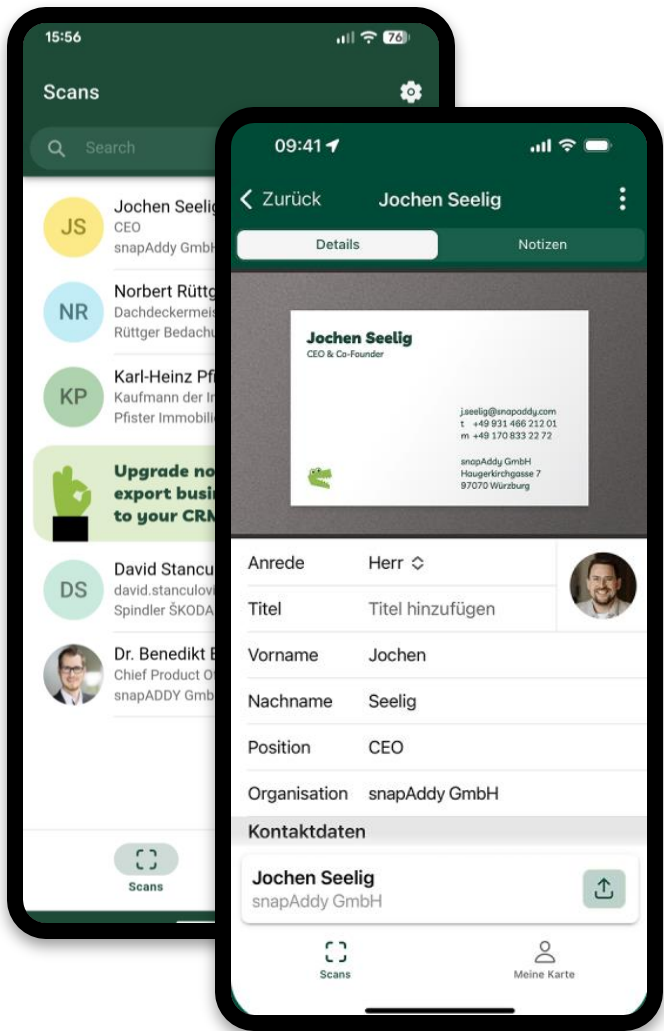
The tool for automated contact
data maintenance in your CRM.

VisitReport

The lead app for trade shows, field
sales, and customer meetings.



Scanning Business Cards.

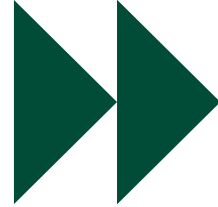
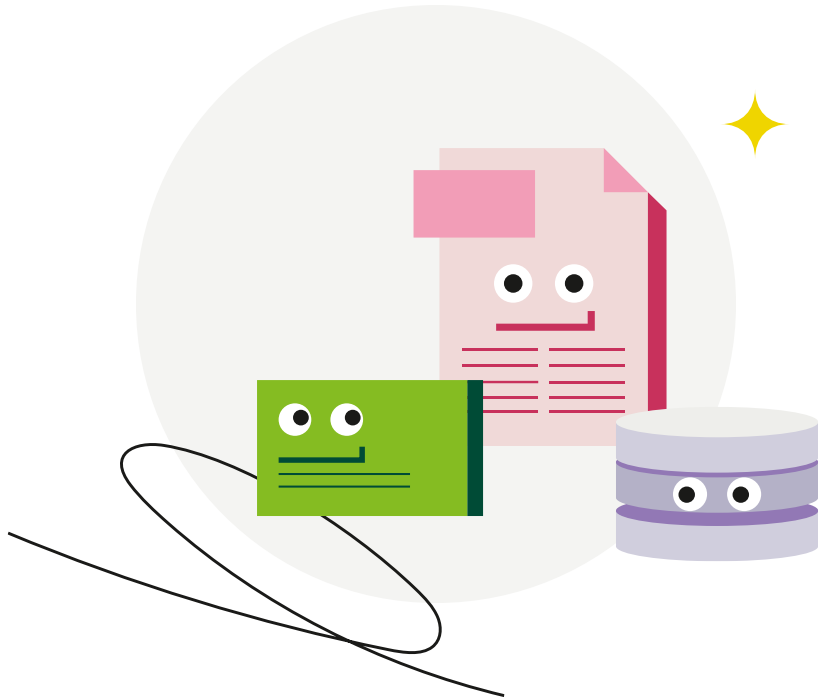


What's the Problem?



Solution:
Fine-tuned deep neural network.

Product Overview.



BusinessCards

Your digital business card – always at hand.

DataQuality


The tool for automated contact data maintenance in your CRM.

VisitReport

The lead app for trade shows, field sales, and customer meetings.



Enriching CRM Contacts with External Data.



Klara Müller

Name

Marketing Manager

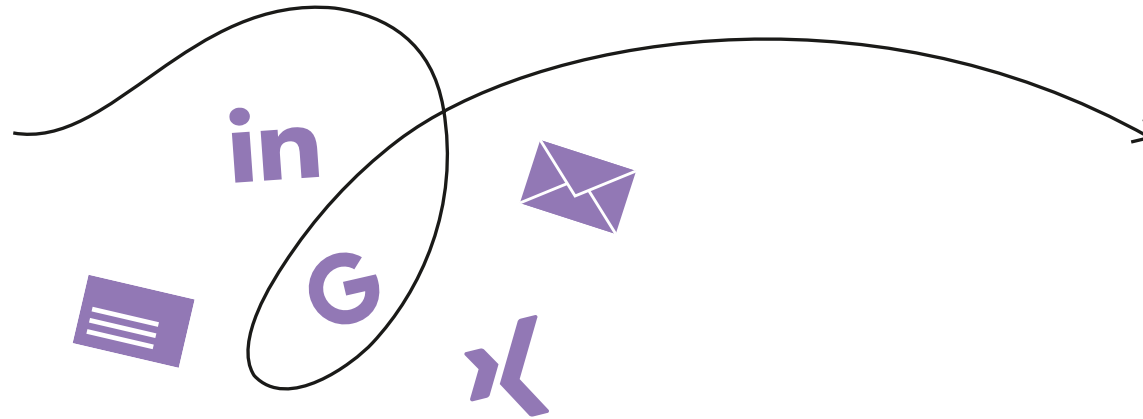
Job title


-

Email

089 999 980 00

Phone





Clara Müller

Name

Head of Marketing

Job title

c.mueller@abc.de

Email

+49 89 999 980 42

Phone

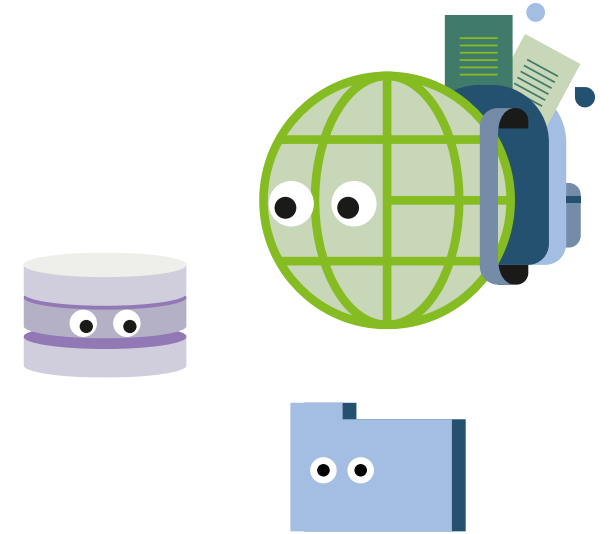


Sources of Interest.

- Company websites (legal notices, imprints, blog posts etc.)
- Official company registries (e.g. **Handelsregister**)
- Google Places / Maps
- Wikipedia
- Signatures of incoming emails
- Public **LinkedIn profiles**
- etc.

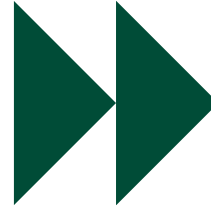
Capstone SS24

Capstone WS25/26



Solution:
**Web scraping and a set of
specific classifiers.**

Product Overview.



BusinessCards

Your digital business card –
always at hand.

DataQuality

The tool for automated contact
data maintenance in your CRM.

VisitReport

The lead app for trade shows, field
sales, and customer meetings.



Voice Assistant for Creating Reports.

Baseline

Users create reports by manually filling in (possibly complex) questionnaires using the VisitReport app.

Voice AI

Let users fill in the questionnaires by talking to a voice assistant instead.



Solution:
A pipeline with several LLMs.

A smartphone mockup of the VisitReport app. The screen displays a questionnaire titled "Company information" with a "DONE" button in the top right. The first section is "Industry" with the prompt "What industry do you work in?" and a red "SELECT" button. The second section is "Company size" with five radio button options: "1-50", "51-100", "101-500", "501-1000", and "1000+". At the bottom of the form are two circular navigation buttons, a grey one with a left arrow and a red one with a right arrow. Below these is a "Talk to Voice AI" button with a microphone icon. The app's bottom navigation bar has three icons: "Reports" (document icon), "Meetings" (calendar icon), and "Media" (photo icon). The status bar at the top shows the time 10:12, a timer at 00:01, and a battery level of 92%.

Section	Field	Options
Company information	Industry	What industry do you work in? [SELECT]
	Company size	<ul style="list-style-type: none">1-5051-100101-500501-10001000+

Data Team.

PDS WS22/23



Severin Simmler
Team Lead Data Science



Wai Lun Wong
Senior Data Scientist



Nele Schneider
Data Scientist



Iryna Musiienko
Data Assistant



Rasha Skeineh
Data Assistant



Dr. Benedikt Brief
Chief Product Officer



Open Jobs.

✓ Working student • Würzburg

**Technical Consultant
System Integration** (m/w/d)

✓ Full-time • Würzburg

Presales Consultant (m/w/d)



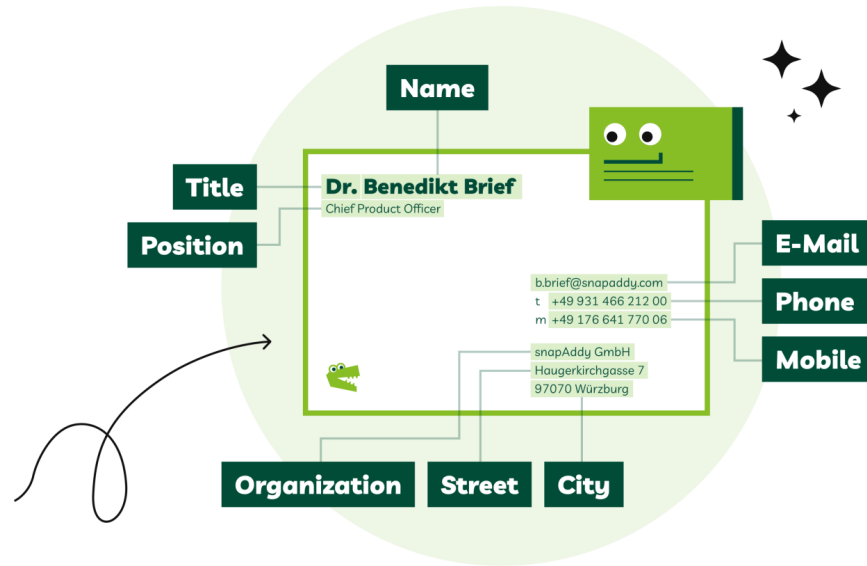
snapAddy.com/jobs

2

AI and Machine Learning at snapAddy.



AI and Machine Learning at snapAddy: Business Card Scanning.

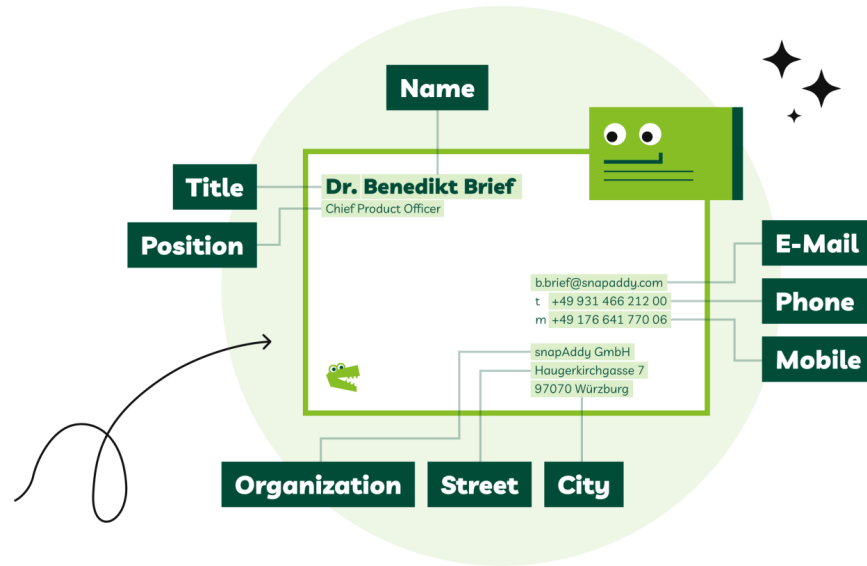


Challenges

- Where to get training data from?
- How to integrate existing domain knowledge?
- Wide variety of possible layouts and formatting
- How to deal with multiple charsets and languages?
- Guarantee consistency and completeness of results?
- Keep inference time below 100ms



AI and Machine Learning at snapAddy: Business Card Scanning.

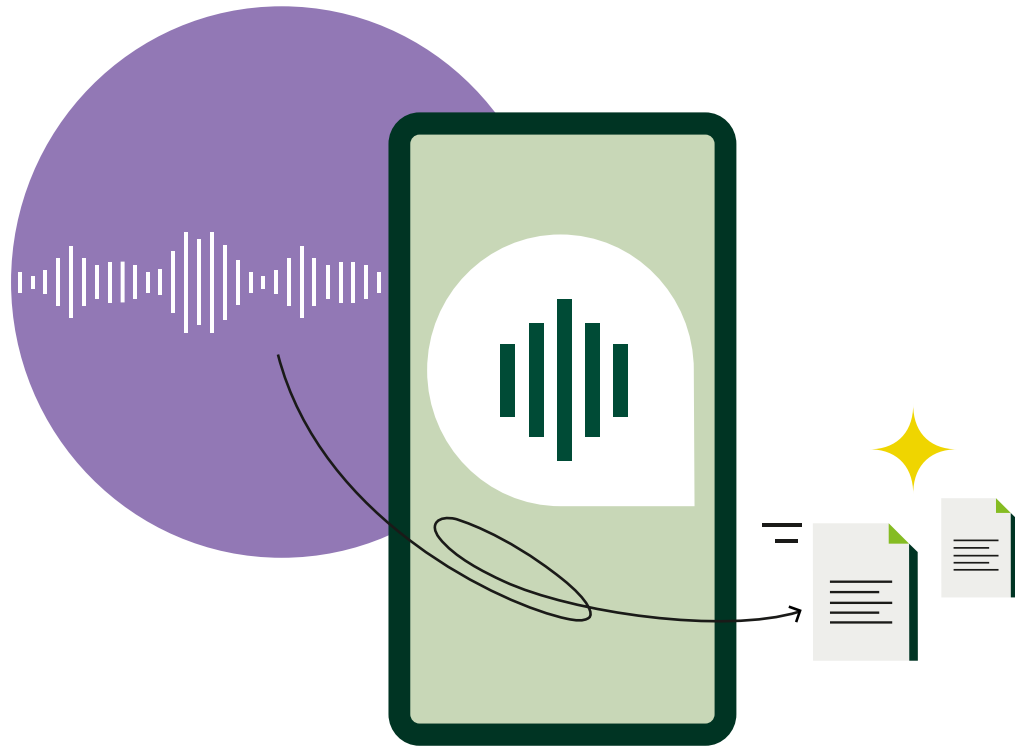


Ideas

- Annotated, high-quality data is key
- Information extraction with regular expressions (email, phone number, postal code etc.)
- Validation of extracted information with a knowledge base (postal code and city etc.)
- Probabilistic named entity recognition for job titles, company names etc.



AI and Machine Learning at snapAddy: Voice AI.

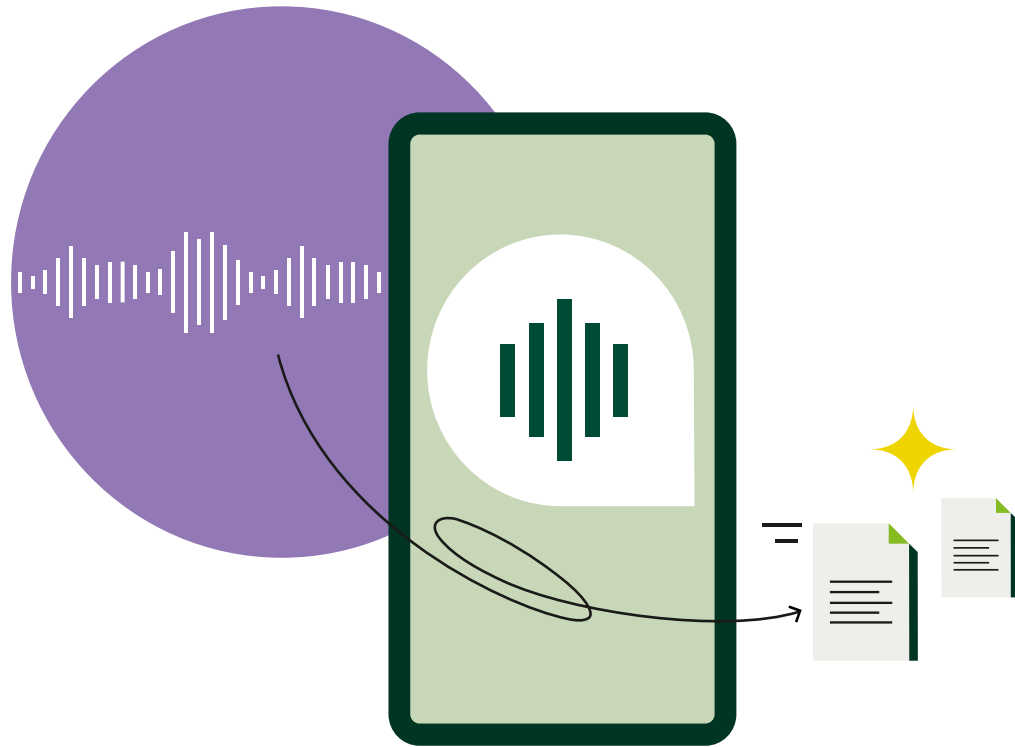


Challenges

- How to turn speech into text?
- What about complex questionnaires?
- Conversation should feel natural, but still strictly focused on the questionnaire
- Avoid hallucinations
- Keep inference time below 100ms?



AI and Machine Learning at snapAddy: Voice AI.



Ideas

- On-device vs. server-side speech recognition
- Structured output vs. natural language
- Summarize unanswered questions to create a natural but focused conversation
- Rule-based plausibility checks to avoid hallucinations
- Unfortunately, little to no control over inference time



3

Don't Believe the Hype: LLMs vs. Traditional Machine Learning.



Don't Believe the Hype.



Pros LLM

- Works pretty good out-of-the-box
- No train and test data necessary
- Quick prototyping

Cons LLM

- Expensive
- Slow
- No train and test data necessary
- Usually external dependency

Pros Traditional Machine Learning

- Works pretty good for a lot of tasks
- Full control over model and architecture
- Fast (depends, of course) and cheap

Cons Traditional Machine Learning

- Not as powerful as LLMs
- Model selection necessary
- Train and test data necessary
- Self-hosted



4

Capstone Project.



Capstone Project.

Your New Lead Research Job

- Your job is to research new leads
- Every week, you receive a list of companies for which you must identify marketing people
- Only senior marketing people and above

How You Would Probably Approach This

- Go to LinkedIn
- Type the first company name and hit the search button
- Start scrolling and identify those profiles with a relevant marketing position





Suche

Start

Ihr Netzwerk

Jobs

Nachrichten

Mitteilungen

Sie

Produkte

[Jetzt Premium für 0 EUR testen](#)

Dein Experte, wenn es um CRM geht.

↓

Jochen Seelig

He/Him · 1.

CEO & Co-Founder der snapAddy GmbH | Wir automatisieren die CRM-Kontaktdatenpflege. | Event & Podcast Host der CRM Experience

Würzburg, Bayern, Deutschland · [Kontaktinfo](#)

[My snapADDY BusinessCard](#)

5.130 Follower:innen · 500+ Kontakte

Fabian Streicher, Mathias Ott und 311 weitere gemeinsame Kontakte

Nachricht

Zur Website

Mehr

Berufserfahrung

CEO & Co-Founder
snapADDY GmbH · Vollzeit
Aug. 2015–Heute · 10 Jahre 5 Monate
Würzburg, Bayern, Deutschland

Senior Consultant SEO/SEA
eology GmbH · Vollzeit
Juli 2013–Juni 2014 · 1 Jahr
Volkach, Bayern, Deutschland

Externer Berater - B2B-Projekte Social Media & Plattformen
Deutsche Telekom · Selbstständig
Juli 2012–Juni 2013 · 1 Jahr
Bonn, Nordrhein-Westfalen, Deutschland

Marketing Manager
Mayflower GmbH · Vollzeit
Jan. 2012–Juni 2012 · 6 Monate
Würzburg, Bayern, Deutschland

Dualer-Student und Projektmanagement
EIKONA AG · Vollzeit
Okt. 2008–Sept. 2011 · 3 Jahre
Volkach, Bayern, Deutschland

Weitere Profile für Sie

Sebastian Metzger · 1.
CTO snapADDY GmbH
TypeScript - Angular - NestJS

Nachricht

Roland Hötzel · 1.
Co-Founder / CFO bei snapADDY GmbH

Nachricht

Matthias Zitzelsberger · 1.
Senior Sales Manager at snapADDY

Nachricht

Victor Weitzmann · 1.
HubSpot Power-User | Making Revenue Teams Play...

Nachricht

Hanna Klüpfel · 1.
Senior Marketing Manager at snapAddy

Nachricht

Alle anzeigen

Kennen Sie ...?

Sandra Brückner
M.Sc. HCI, UX-Manager:in bei HMI Project

Vernetzen

Jan Ludwig
M. Sc. Digital Business Systems Student | QA @ snapAddy | D...

Vernetzen

Marie Louise Eichner
In Ausbildung/Studium: Julius-Maximilians-Universität...

Capstone Project.

Your New Lead Research Job

- Your job is to research new leads
- Every week, you receive a list of companies for which you must identify marketing people
- Only senior marketing people and above

How You Could Optimize This

- Automatically search for candidates for each company
- Identify the candidates' department and seniority based on their job title
- Filter out candidates who are not working in marketing and who are not at least a senior
- Manually review the candidates left and identify relevant profiles



Capstone Project.

Your New Lead Research Job

- Your job is to research new leads
- Every week, you receive a list of companies for which you must identify marketing people
- Only senior marketing people and above

How You Could Optimize This

- Automatically search for candidates for each company
- **Identify the candidates' department and seniority based on their job title**
- Filter out candidates who are not working in marketing and who are not at least a senior
- Manually review the candidates left and identify relevant profiles



Capstone Project.

The Data

- 609 CVs from real-world public LinkedIn profiles with 2.637 job titles and labels for department and seniority
- 390 additional CVs with 1.886 job titles without labels

Additional Data

- 9.428 job titles with seniority labels
- 10.145 job titles with department labels

The Goal

```
>>> pipeline("CMO")  
{ "department": "Marketing", "seniority": "Management" }
```



Capstone Project.

Departments

- Marketing
- Sales
- Information Technology
- Business Development
- Project Management
- Consulting
- Administrative
- Purchasing
- Customer Support
- Human Resources
- Other

Seniority

1. Management (e.g., Vice President, Geschäftsführer)
2. Director (e.g., Marketing Director)
3. Lead (e.g., Head of IT, Team Lead Marketing)
4. Senior (e.g., Senior Software Engineer)
5. Professional (e.g., Software Engineer)
6. Junior (e.g., Junior Software Engineer)



Bye!

