



# ANNA GUNAZA

## IT CONSULTANT

### KONTAKT

- ☎ 015678134406
- ✉ hello@schroeck-consulting.de
- 🌐 [www.schroeck-consulting.de](http://www.schroeck-consulting.de)
- 📍 Leipziger Platz 12, 10117 Berlin

### AUSBILDUNG

**Mathematik Universität:**  
National University of Kyiv,  
Ukraine  
**Zeitraum:** 2020 – 2024  
**Abschluss:** Bachelor of Science

**Mathematik -**  
**Auslandssemester**  
**Universität:** Freie Universität  
Berlin, Deutschland  
**Zeitraum:** 2023

### SPRACHEN

**Deutsch:** fließend  
**Englisch:** fließend  
**Ukrainisch:** Muttersprache

### ZERTIFIZIERUNGEN

**Databricks**  
Certified Data Engineer  
Associate, 2025

### PROFIL

*“Als Mathematikerin mit Erfahrung im Energiesektor arbeite ich als Data Analyst mit Fokus auf Data Engineering, Analytics, Machine Learning und Künstliche Intelligenz (KI). Ich nutze Python, Statistik und Datenbanken, um interaktive Dashboards und datenbasierte Lösungen zu entwickeln. Mit Teamgeist und Engagement treibe ich innovative Projekte voran und erweitere kontinuierlich meine analytischen Fähigkeiten“*

### SKILLS

#### Programmiersprachen & Frameworks

- **Python:** Pandas, NumPy, Plotly, Scikit-learn, TensorFlow, Matplotlib, Seaborn, pytest, unittest, SciPy
- **Java**
- Streamlit, Flask, Django, FastAPI, Dash

#### Datenanalyse & -verarbeitung

- **Machine Learning:** Scikit-learn, TensorFlow
- **Natural Language Processing (NLP):** GPT/Generative AI, Chatbot-Entwicklung
- **Data Engineering & Data Modelling:** Databricks, dbt, Snowflake
- **Dashboard-Erstellung:** Power BI, Jupyter Notebook, Matplotlib, Plotly, Seaborn
- **SQL**

#### Tools & Plattformen

- **Version Control:** GitHub, Gitlab, Azure DevOps
- **CI/CD & Cloud:** Azure DevOps, Snowflake
- **Workflow & Projektmanagement:** Jira, Confluence, Scrum
- **APIs & Sicherheitsprotokolle**
- **OAuth**

## PROJEKTE

---

### **Python AI Developer**

Medienunternehmen Berlin | Januar 2025 – Laufend

#### **Entwicklung einer Prozessautomatisierung mittels LLM's in HR**

Entwicklung eines Streamlit-basierten Tools zum automatischen Ausfüllen von PDFs auf der Grundlage von Mitarbeiterstammdaten. Das Tool verwendet OpenAI OCR, um automatisch ausfüllbare Felder zu erkennen und Mitarbeiterstammdaten an die richtigen Stellen zu schreiben. Die Lösung senkte die Aufwände für die Dokumentenverwaltung um 50 %.

Anna hat das komplette Projekt betreut, inklusive Anforderungs-Workshops, Aufgaben- und Featureplanung über die Umsetzung bis zur Integration in die Systemlandschaft des Kunden.

**Technologien:** Vektordatenbanken, Langchain, OpenAI/ChatGPT, Python, Streamlit, Pandas, Rest API, SQL Datenbank, OCR, Maschinelles Lernen

---

### **Data Engineer / ETL-Entwickler Energiehandel**

Energieversorger, Essen | Mai 2024 – Dezember 2024

Entwicklung von Datenanalysen und Datenanalysen-Tools für Energie-Trader.

Aufbau von ETL-Strecken von ENTSOE, BMRS, Copernicus, Volue und anderen mittels Databricks.

Integration der Daten in Snowflake, Visualisierung in PowerBI und Streamlit.

Anna hat zudem komplexe Analysen gebaut, die den Tradern helfen, fundierte Entscheidungen und Vorhersagen zu treffen.

Insbesondere für folgende Daten:

- Wettervorhersage-Daten, da in der Energiewende sehr relevant zur Vorhersage von Preisen
- Live-Daten Energie-Markt, Vorhersagen und historische Daten
- Verfügbarkeiten und Residual-Last-Daten geladen, visualisiert und transformiert gemäß Anforderungen der Analysten

**Skills:** JIRA, Streamlit, Data Integration, Snowflake, Datenanalyse, Databricks, PowerBI, Volue, KPLER

---

### **Python AI Developer**

Kaffehändler, Berlin | Oktober 2024 – Januar 2025

#### **Entwicklung einer Dateisuche in ELO mittels GPT**

Entwickelte ein Streamlit-basiertes Tool zur Optimierung der Dokumentsuche innerhalb von Ordnerstrukturen in ELO.

Dieses Tool adressiert die häufig ineffiziente und unzuverlässige integrierte Textsuche in Unternehmen. Es ist sowohl in ELO als auch in anderen Lösungen einsetzbar und kann sogar lokal auf dem eigenen Rechner verwendet werden. Das Tool nutzt fortschrittliche Technologien wie Vektordatenbanken und generative Sprachmodelle (LLM/GPT), um relevante Dokumente basierend auf Kontext oder spezifischen Textausschnitten zu identifizieren. Dadurch wird die Suche nach wichtigen Informationen erheblich beschleunigt und verbessert.

**Technologien:** Vektordatenbanken, ELO, OpenAI/ChatGPT, Python, Streamlit, Pandas, Rest API

---

### **Python AI (GPT) Developer – Entwicklung einer Dokumenten-Validierungs-App**

Immobilienfinanzierer, Berlin | Juni 2024 – September 2024 | 3 Monate

Entwicklung einer benutzerfreundlichen Streamlit-App zur Dokumentenvalidierung im Immobilienfinanzierungsprozess, die Upload, Prüfung und sichere Verarbeitung zentraler Dokumente unterstützt.

- **App zur Dokumentenverarbeitung:** Entwicklung einer Streamlit-App, die den Upload und die Verwaltung relevanter Dokumente (z. B. Ausweise, Grundbuch-Einträge, Kontoauszüge) im Finanzierungsprozess vereinfacht.
- **KI-gestützte Validierung:** Integration von ChatGPT zur automatisierten Validierung und Überprüfung hochgeladener Dokumente, wodurch die Effizienz und Genauigkeit des Prüfprozesses verbessert wurden.
- **Datenschutzkonzept:** Entwicklung eines umfassenden Datenschutzkonzepts zur Sicherstellung der Einhaltung regulatorischer Anforderungen und zur Gewährleistung der Datensicherheit.
- **Mitarberschulung:** Onboarding und Schulung der Mitarbeitenden zur sicheren und effektiven Nutzung der Anwendung.

**Skills:** Vektordatenbanken, Python, Streamlit, GPT, Prompt Engineering, Prozessentwicklung und -optimierung

---

### **Data Analyst/Data Scientist – Kundenzufriedenheitsanalyse mit NLP und Dashboards**

Energieversorger, München und Essen | 2023-2024 | 1 Jahr 3 Monate

In meiner Rolle als Data Analyst war ich maßgeblich am Net Promoter Score (NPS)-Projekt beteiligt, um datenbasierte Einblicke in die Kundenzufriedenheit zu gewinnen und das Unternehmen strategisch zu unterstützen. Durch den Einsatz von Natural Language Processing (NLP) und Machine Learning konnte ich wertvolle Erkenntnisse aus Kundenfeedback gewinnen und dazu beitragen, dass datenbasierte Optimierungen identifiziert und priorisiert wurden.

Hauptaufgaben und Erfolge:

- **Sentiment-Analyse zur Kundenzufriedenheit:** Durchführung detaillierter Sentiment-Analysen, um das Stimmungsbild der Kunden zu erfassen und gezielte Verbesserungsmaßnahmen zu identifizieren.

- **Topic Modeling für thematische Einblicke:** Identifikation von Hauptthemen im Kundenfeedback mittels Topic Modeling, um die Kundenerfahrungen tiefgehend zu verstehen und gezielte Optimierungsvorschläge abzuleiten.
- **Aufbau eines datenbasierten Konzepts und Dashboard-Integration:** Beginnend mit einer strukturierten Datenorganisation und Voranalyse entwickelte ich ein datenbasiertes Konzept, das relevante KPIs ableitete und diese in interaktive Dashboards integrierte. So wurde eine kontinuierliche Überwachung und datenbasierte Entscheidungsfindung im Management ermöglicht.
- **Experimentierung mit Large Language Models (LLMs):** Einsatz moderner Sprachmodelle zur Verbesserung der Analysegenauigkeit und Tiefe der Erkenntnisse.
- **Effiziente Datenverarbeitung und Visualisierung:** Nutzung von Jupyter Notebook zur interaktiven Datenexploration und Visualisierung sowie Databricks für skalierbare Datenverarbeitung und anspruchsvolle Analysen.
- **Versionierung und Zusammenarbeit mit GitLab:** Organisierte Zusammenarbeit und Versionskontrolle durch GitLab und Automatisierung der Analysepipelines, um zuverlässige und reproduzierbare Ergebnisse sicherzustellen.

**Skills:** Machine Learning, Natural Language Processing, Aufbau von Datenkonzepten, Dashboard-Erstellung, Large Language Models, Python, Jupyter Notebook, Microsoft Excel, GitLab, GitLab Pipelines, Databricks, pytest, unittest

---

### **Python Software Engineer – Entwicklung und Optimierung eines Bonusberechnungssystems** Energieversorger, München und Essen | 2022-2023 | 6 Monate

In diesem Projekt entwickelte ich eine robuste, Python-basierte Webanwendung mit Flask und FastAPI zur Automatisierung und Überwachung des Bonusberechnungsprozesses für alle Vertriebskanäle. Der Prozess, der bisher vier Stunden in Anspruch nahm, wurde durch meine Optimierungen deutlich effizienter und zuverlässiger gestaltet.

Hauptaufgaben und Erfolge:

- **Leistungsoptimierung durch Caching:** Implementierung eines Caching-Systems, das die Antwortzeiten der Anwendung und die Effizienz der Bonusberechnung signifikant verbesserte.
- **Sicherheitsinfrastruktur mit OAuth-Authentifizierung:** Integration von OAuth via Entra ID, um eine sichere und kontrollierte Anwendung für den Zugriff durch Vertriebsteams zu gewährleisten.
- **CI/CD-Pipeline und DevOps-Verwaltung:** Aufbau und Verwaltung einer DevOps-Pipeline zur nahtlosen Bereitstellung und Versionierung der Anwendung, inklusive Testing und Monitoring mit pytest und unittest.
- **Zusammenarbeit mit dem Datenteam:** Enger Austausch mit dem Datenteam zur Sicherstellung der optimalen Datenstrukturierung und -verfügbarkeit in Snowflake, was die Genauigkeit und Konsistenz der Bonusberechnungen stärkte.

**Skills:** Python, SQL, Data Modelling, Flask, FastAPI, Git, Snowflake, Azure DevOps Pipelines, EntraID, pytest, unittest

---

## Data Analyst – Textanalyse und Dashboard-Erstellung für Telegram-Chats

Freelance | Kiew | 2022 | 6 Monate

In diesem Projekt führte ich eine detaillierte Datenanalyse von Kommunikationsaktivitäten und -mustern in Telegram-Chats durch. Mithilfe von Python und spezialisierten Bibliotheken zur Datenverarbeitung und -visualisierung erstellte ich interaktive Dashboards, die datenbasierte Einblicke in das Nutzerverhalten und die Plattform-Interaktionen ermöglichten.

Schwerpunkte der Analyse und Dashboard-Erstellung:

- **Gesamtaktivität:** Darstellung der Nachrichtenhäufigkeit und -verteilung in einem Dashboard, um aktive Nutzungsphasen und Kommunikationsspitzen zu identifizieren.
- **Nachrichtenverteilung:** Visualisierung von Sende- und Empfangsverhältnissen in einem interaktiven Dashboard, um Interaktionsmuster zwischen Nutzern und Gruppen sichtbar zu machen.
- **Nachrichtentypen:** Erstellung einer Kategorie-Analyse (Text, Bild, Video), die Kommunikationspräferenzen und den Informationsfluss verdeutlicht.
- **Durchschnittliche Nachrichtenlänge:** Analyse und Visualisierung der Textlänge in Dashboards zur Bewertung von Kommunikationsstilen und Interaktionsintensität.
- **Geschlechterverteilung:** Demografische Analyse und Dashboard zur Darstellung der Geschlechterverteilung in verschiedenen Gruppen, um spezifische Nutzungsmuster zu identifizieren. Für alle Key Features wurden Visualisierungen erstellt, um die Erkenntnisse aus den Daten effektiv zu vermitteln. Ich habe Python und verschiedene Datenverarbeitungs- und Visualisierungsbibliotheken verwendet, z. B. Pandas, Matplotlib, Seaborn und Plotly.

**Skills:** Python, Pandas, Matplotlib, Seaborn, Plotly, Dashboards, Datenvisualisierung, Textanalyse

---

## Software Engineer – Entwicklung eines automatisierten Lagerverwaltungssystems

Freelance, Kiew | 2021 | 6 Monate

Im Rahmen eines Projekts zur digitalen Transformation eines kleinen Unternehmens habe ich ein maßgeschneidertes Automatisierungssystem für die Lagerverwaltung entwickelt. Ziel war es, manuelle Prozesse zu eliminieren und die Bestandsführung effizient und fehlerfrei zu gestalten.

Highlights der Implementierung:

- **Benutzerfreundliche GUI:** Entwicklung einer intuitiven grafischen Oberfläche mit Swing, die eine einfache Bedienung und schnelle Einarbeitung für Lagerpersonal ermöglicht.
- **Automatisierte Bestandsverwaltung:** Funktionen zur präzisen Erfassung von Produktgruppen und Artikeln, einschließlich Hinzufügen, Bearbeiten und Löschen, um die Datenqualität zu verbessern.

- **Transaktions-Tracking und Reporting:** Einführung einer automatisierten Bestandsaufzeichnung und detaillierten Berichte zur Transparenz von Lagerbeständen, die datenbasierte Entscheidungen für das Management unterstützen.
- **Effiziente Suche und Filterung:** Implementierung von Suchfunktionen, um Produktinformationen in Echtzeit bereitzustellen und die Lagerabläufe zu beschleunigen.

Dieses Projekt stellte hohe Anforderungen an **Java** und **UI/UX-Design mit Swing** sowie an die **Datenverwaltung und Prozessautomatisierung**. Meine Fähigkeit, Geschäftsanforderungen in technische Lösungen zu übersetzen, ermöglichte eine deutliche Effizienzsteigerung im Lager und verbesserte die Datenzugänglichkeit für das Management.

**Skills** Java, Swing, UI/UX Design, Prozessautomatisierung, Datenanalyse, Bestandsoptimierung