



netidee

PROJEKTE

Zwischenbericht | Call 19 | Projekt ID 7377

Lizenz: CC BY / CC BY-SA

1 Einleitung

Das ist der Zwischenbericht für das Projekt „Leichte Sprache KI“, das zum Ziel hat, ein kostenloses Tool ins Leben zu rufen, das Texte in “Leichte Sprache” übersetzen kann.

Das Projekt umfasst eine kostenlose API und ein Tool, das auf eigenen Servern betrieben werden kann. Viele Menschen haben Schwierigkeiten, komplexe Texte zu verstehen, was ihre Teilhabe einschränkt. Unser Tool verbessert die Verständlichkeit von Informationen und fördert so die Inklusion und Barrierefreiheit für Menschen mit Lernschwierigkeiten, Nicht-Muttersprachler und andere Betroffene.

2 Status der Arbeitspakete

2.1 AP1: Detailplanung und Formales am Projektstart

Dieses Arbeitspaket umfasst die folgenden Punkte:

- Vertrag unterschrieben, Detailprojektplan (Arbeitsblatt Arbeitspakete) erstellt und abgenommen,
- detaillierte Liste Projektergebnisse mit Lizenz und Ort der öffentlichen Bereitstellung erstellt und abgenommen (Arbeitsblatt Projektergebnisse)
- Projekt-Website in Betrieb & erster Blogeintrag erstellt
- erste Förderrate beantragt; ggf. nach Feedback ergänzt
- Förderrate1 genehmigt

Das Arbeitspaket wurde vollständig umgesetzt. Insgesamt wurden 2 Stunden mehr gebraucht als veranschlagt.

2.2 Arbeitspaket 2 - Konzeption

Dieses Arbeitspaket umfasst die folgenden Punkte:

- Workshop(s) zum Thema Leichte Sprache mit ExpertInnen
- Konzeption für die Ausgestaltung des Sprachmodells
- Konzeption für die technische Umsetzung und des Hosting-Systems
- Ausarbeitung Business Modell

Projektergebnis: Workshop zu Leichte Sprache

Gemeinsam mit unserem Partner Okay - Büro für Leichte Sprache haben wir einen Workshop durchgeführt, um intern Know How aufzubauen, aus der Praxis zu lernen und die Basis für das technische Konzept zu schaffen.

Einen Tag lang haben wir in großer Runde mit Helga Mock, einer Expertin für Leichte Sprache, die Grundlagen besprochen und Praxisbeispiele analysiert.



Abb.: Foto vom Workshop mit Helga Mock

Durch die Expertise von Helga Mock konnte das Team grundlegende Konzepte erfahren, Probleme aus der Praxis bei der Arbeit mit Leichter Sprache kennenlernen und wichtige Insights erfahren, die für die Konzeption des Systems von Bedeutung waren.

Projektergebnis: Konzeption für die Ausgestaltung des Sprachmodells (Technisches Konzept)

Die Konzeption selber gestaltete sich schwieriger als erwartet. Je mehr wir in die Materie eintauchten und konzipierten, desto mehr stellte sich heraus, dass ein einstufiger Prozess, wie ursprünglich vorgesehen, unseren Anforderungen bzgl. Qualität nicht gerecht werden kann. Gemeinsam mit dem Fehlen von verwendbaren Trainingsdaten, stellte uns diese Erkenntnis vor Herausforderungen. Wir haben uns dann nach einigen Diskussionsrunden für die Umsetzung eines mehrstufigen Prozesses entschieden. Das bedeutet konkret, dass nicht nur ein LLM Modell trainiert werden muss, sondern mehrere Modelle und Agenten bei der Übersetzung in Leichte Sprache zusammenarbeiten.

Also Konsequenz waren jedoch ein paar personelle Änderungen notwendig. Wie in der Einleitung beschrieben, übernahm Markus Müller die externe Beratung bzgl. Konzeption.

Zur Entscheidung bzgl. der Prozessarchitektur haben wir einen Blogpost veröffentlicht, und Details dazu sind im Technischen Konzept einsehbar (verfügbar als eigenes Dokument als Projekt-Ergebnis).

Projektergebnis: Konzeption des Hosting Systems

Das Hosting-Konzept basiert auf einer kosteneffizienten, skalierbaren Lösung mit folgenden Kernmerkmalen:

Hardware:

- Intel Arc Pro B60 GPU (Blower-Design) als performante und kostengünstige Alternative zu NVIDIA
- Intel Xeon Server mit 64 GB RAM und ausreichend Storage
- Optimierte für AI-Inferenz-Workloads

Standort:

- Österreichisches Rechenzentrum (2 Optionen in Evaluation)
- DSGVO-konform und optimal für DACH-Zielgruppe

Kosten:

- Deutliche Ersparnis gegenüber Cloud-GPU-Lösungen

Skalierbarkeit:

- Start mit Single-Server-Setup
- Klare Upgrade-Pfade bei Wachstum
- Hardware-Erweiterungen ohne Architektur-Änderungen möglich

Diese Infrastruktur bietet eine solide Basis für den Produktstart mit der Flexibilität, bei Erfolg entsprechend zu wachsen. Mehr Details dazu im Hosting-Konzept (verfügbar als eigenes Dokument als Projekt-Ergebnis).

Projektergebnis: Businessplan

LeichteSprach KI ist ein Open-Source-Projekt, das durch

- eine netidee-Förderung die technische Grundlage erhält.
- Durch Sponsoring langfristige Finanzierung sicherstellt,
- mit Custom Training wesentliche Einnahmen generiert,
- von der Community gemeinsam verbessert und weiterentwickelt wird
- und Transparenz Vertrauen in sensible Technologie schafft.

Die Finanzierungsstrategie besteht aus folgenden Punkten:

- Keine externe Finanzierung notwendig
- Diversifizierte Einnahmequellen (Sponsoring, Custom Training, Förderungen)
- Profitabilität ab 2027
- Langfristig selbsttragend

Mehr dazu im Business Plan Dokument (verfügbar als eigenes Dokument als Projekt-Ergebnis).

2.3 Arbeitspaket 3 - Programmierung und Hosting Basis-System

Dieses Arbeitspaket umfasst die folgenden Punkte:

- Einrichtung Entwicklungsumgebung
- Programmierung Prototyp für Basis-System
- Ausprogrammierung Basis-System
- Programmierung Software für die Bereitstellung der API

Ergebnis: Entwicklungsumgebung

Die Entwicklungsumgebung wurde eingerichtet und ist im Einsatz.

Ergebnis: APP Entwicklung

Hier haben wir einige Systeme ausprobiert und haben einen rudimentären Prototyp mit <https://tauri.app/> umgesetzt. Vorteil ist, dass wir damit wahrscheinlich alle gängigen Systeme abbilden können. Linux, macOS, Windows, Android und iOS - alles mit der gleichen Codebase. Es muss sich aber erst zeigen, ob wir diesen Weg gehen können, da der Aufwand viel höher ist, um mehrere Systeme zu unterstützen.

Ergebnis: Programmierung Prototyp für Basis-System

Ein zentrales Tool hierzu ist die Leichte Sprache Checker API, die bereits verfügbar ist und weiter optimiert wird.

Im Zuge der Recherche und Tests sind wir zur Erkenntnis gelangt, dass wir ein Service benötigen, welches die Qualität von Leichte Sprache Texte analysiert und quantifiziert. Dieses Tool hatten wir vorab nicht angedacht, aber ist nach Recherche und ersten Tests maßgeblich für das weitere Vorgehen bei dem Projekt.

Nur wenn wir Texte verstehen können, können wir diese auch erstellen. Und wir können so auf die fehlenden Trainingsdaten eher verzichten.

Dieses Service nennen wir “Leichte Sprache Checker” und besteht aus vielen Regeln, die jeweils ein eigenes Modell verwenden. Das kann ein vorhandenes Open Source Modell sein oder ein von uns neu trainiertes. Wir haben einen Blogpost veröffentlicht, der Details dazu beinhaltet und den aktuellen Status beschreibt.

Die Checker API wird wesentlicher Teil des Basis-Systems sein. Für einen vollen Prototypen fehlen dann noch die Prozessschritte für die Extrahierung der Kerninformation und die Einbindung der

Evaluierungsergebnisse für die Optimierung der Ergebnisse – das ist aktuell in Arbeit. Der Aufwand insgesamt für dieses Arbeitspaket wird größer sein, als ursprünglich angenommen. Der in der Konzeption eingeschlagene Weg, einen mehrstufigen Prozess abzubilden, erhöht den Programmieraufwand (statt einem Modell werden mindestens 3 zum Einsatz kommen, im Leichte Sprache Checker selber noch mal einige mehr).

Als Konsequenz wird sich der Fertigstellungstermin für dieses Arbeitspaket um 4 Monate nach hinten verschieben. Neben dem erhöhten Arbeitsaufwand wurden auch personelle Änderungen nötig. Dazu kommt, dass im Juli und August krankheits- und urlaubsbedingt wenig am Projekt gearbeitet werden konnte.

Wir sind jedoch zuversichtlich, den neue Zeitplan auch mit den hohen uns selber gestellten Anforderungen bzgl. Qualität einhalten zu können.

2.4 Arbeitspaket 4 - Programmierung Browser Plugin

Dieses Arbeitspaket umfasst die folgenden Punkte:

- Programmierung Browser Plugin für Chrome
- Programmierung Browser Plugin für Firefox (optional)

Erreichte Ergebnisse

Erster Prototyp für das Browser Plugin ist vorhanden (Ohne Funktion, aber als vordefiniertes WebExtensions API, welche von Chrome, Firefox und mit Apple-spezifischen Einschränkungen auch in Safari funktioniert).

2.5 Arbeitspaket 5 - Marketing und PR

Dieses Arbeitspaket umfasst die folgenden Punkte:

- Erstellung Website für Leichte Sprache inklusive Darstellung und Präsentation des Projektes
- Erstellung Dokument für Externkommunikation zur Veröffentlichung auf der netidee Website
- Konzeption, Produktion eines Werbevideos
- Erstellung von Design-Guidelines und Dokumentation
- Erstellung von Graphiken und Icons für das Projekt
- Erstellung Logo für das Projekt
- Präsenz auf Konferenz(en) und Vermarktung des Produktes

Projektergebnis: Definition des USP

Folgende Alleinstellungsmerkmale haben wir identifiziert. Sie sind relevant für das Marketing und Grundlage für die Kommunikation nach außen:

1. 100% Open Source - Volle Transparenz, Community-getrieben
2. Kostenlos für alle privaten Nutzer - Keine Bezahlschranken (aber Rate Limits)
3. Made in Austria - DSGVO-konform, Intel Arc, lokale Infrastruktur
4. Wissenschaftlich fundiert - Regelwerk-basiert, reproduzierbar
5. Selbst-hostbar - Volle Datenkontrolle für Organisationen
6. Community-First - Entscheidungen transparent, partizipativ

Projektergebnis: Logo

Logo Konzept wurde erstellt und das Logo entwickelt:



Sprechblase als Basis, die schwebt.

Die Schriftart ist eine extra auf Leichter Lesen optimierte Font.

Projektergebnis: Erstellung Website für Leichte Sprache

Die Website ist in Arbeit und wird unter der Domain leichte-sprache.ai veröffentlicht werden.

2.6 Arbeitspaket 6 – Projektmanagement

Dieses Arbeitspaket umfasst die folgenden Punkte:

- Projektplanung
- Erstellung Projekt in Jira
- Erstellung Sprints und Projekt-Meetings
- Erfolgskontrolle
- Testing

Das Projektmanagement ist aufgesetzt und begleitet das Projekt wie geplant. Projektplanung ist grundsätzlich abgeschlossen bzw. wird laufend angepasst. Wir verwenden Jira für die Planung und Zusammenarbeit.

3 Zusammenfassung Planaktualisierung

Die Konzeptphase ist abgeschlossen, wir arbeiten aktuell an der Finalisierung des Prototypen für ein Tool, das die Qualität von Leichter Sprache Texte erkennen kann. Dieses Tool ist dann die Grundlage für das Training eines LLM Models.

Das Hauptproblem, das bei der technischen Konzeption deutlich wurde, war der Mangel an Trainingsdaten. Wir haben keinen Zugang zu einer so großen Anzahl von qualitativ hochwertigen Texten in Leichter Sprache, die für das Training eines LLM Models nötig ist. Deswegen haben wir uns dafür entschieden, zuerst ein Tool zu entwickeln, das die Qualität von Texten bewerten kann (mehrstufiger Prozess). Damit sollte es dann möglich sein, auch mit wenigen Trainingsdaten gute Ergebnisse zu erzielen, indem Feedback über die Qualität der Ergebnisse beim Finetuning des Models verwendet werden kann. Nur so können die Ergebnisse wirklich von guter Qualität sein und den Anforderungen gerecht werden.

Das bedingte auch kleine Änderungen bei den Personen, die die Umsetzung durchführen. Markus Müller übernahm die externe Betreuung und Beratung bzgl. der technischen Konzeption. Dafür werden wir bei der Programmierung mehr auf interne Ressourcen setzen können. Externe Expertise werden wir besonders noch für das Training bzw. Finetuning des LLM Models benötigen.

Der voraussichtliche Fertigstellungstermin für die geplanten Ergebnisse und den Endbericht wird sich aufgrund der technischen Komplexität der gewählten Umsetzungsstrategie voraussichtlich auf April 2026 verschieben.

Insgesamt wird sich der Fertigstellungstermin für das Projekt voraussichtlich um 2 Monate nach hinten verschieben. Wir hoffen, dass diese Änderung des Zeitplans möglich ist und freigegeben wird.

	Startdatum	Enddatum	
AP 1: Detailplanung	01.03.2025	12.06.2025	Abgeschlossen
AP 2: Konzeption	01.03.2025	30.06.2025	Abgeschlossen
AP 3: Programmierung	01.07.2025	28.02.2026	Verlängert um 4 Monate
AP 4: Browser Plugin	01.11.2025	28.02.2026	Verlängert um 2 Monate
AP 5: Marketing und PR	01.05.2025	30.04.2026	Verlängert um 2 Monate
AP 6: Projektmanagement	01.03.2025	30.04.2026	Verlängert um 2 Monate
AP 7: Dokumentation und Formales	01.11.2025	30.04.2026	Verlängert um 2 Monate

4 Öffentlichkeitsarbeit/ Vernetzung

Unser Vision für 2028 bzgl. Social Impact:

- 300.000+ Menschen nutzen barrierefreie Texte
- 1 Mio.+ Texte übersetzt
- 30+ Partner-Organisationen
- 50+ Open-Source-Contributors
- Marktstandard für Open-Source Leichte-Sprache-Tools

Wie wollen wir das erreichen?

- Kooperationen mit Bildungseinrichtungen und NGOs: Zusammenarbeit mit Schulen, Universitäten und sozialen Einrichtungen, um das Tool in ihre täglichen Abläufe zu integrieren.
- Partnerschaften: Kooperation um das Tool in bestehende und neue Anwendungen zu integrieren. Z.b News-Medien, für die wir schon Dienstleister in Web Angelegenheiten sind oder potentielle Partnerschaft
- Open Source Community-Beteiligung durch Entwicklungssprints und Implementierungs-Sprints.
- Aufbau und Pflege einer aktiven Entwickler-Community, um partizipativ daran teilzunehmen (Forum, Mailingliste, regelmäßige Meetings.
- Intensive Breite Öffentlichkeitsarbeit durch Social Media (hier komplexer Text -> Hier einfach verständliches Ergebnis), Blogs, Webinare
- Teilnahme an Konferenzen (Open Source Summit und FOSDEM) , um das Bewusstsein für das Tool und seine Vorteile zu erhöhen.
- Medienkampagnen: Gezielte Medienkampagnen, um das Tool in der Öffentlichkeit bekannt zu machen und seine Nutzung zu fördern.
- Implementierung von Erweiterten Feedback-Mechanismen, um kontinuierlich Rückmeldungen von Nutzern zu sammeln für Verbesserungen.
- Regelmäßige Evaluationsprogramme zur Bewertung der Effektivität und des Nutzens des Tools, um den langfristigen Impact zu messen.
- Nachhaltige Finanzierung durch Förderanträge und Sponsoring zur Sicherstellung der nachhaltigen Finanzierung und Suche nach Sponsoren und Partnern.