



netidee

PROJEKTE

PrototAlp

Zwischenbericht | Call 16 | Projekt ID 5899

Lizenz: CC-BY

Inhalt

1	Einleitung.....	3
2	Status der Arbeitspakete.....	3
2.1	Arbeitspaket 1 - Detailplanung und Formales am Projektstart.....	3
2.2	Arbeitspaket 2 - Use-case-Evaluierung.....	3
2.3	Arbeitspaket 3 - Setup und Management der Infrastruktur	4
2.4	Arbeitspaket 4 - Software-Setup	5
3	Umsetzung Förderauflagen.....	5
4	Zusammenfassung Planaktualisierung	6
5	Öffentlichkeitsarbeit/ Vernetzung.....	6
6	Eigene Projektwebsite.....	6

1 Einleitung

Das Projekt “PrototAlp” hat zum Ziel, eine Plattform für die Durchführung von Data-Science-Experimenten, Entwicklung von Machine-Learning-Modellen, sowie dem Experimentieren mit neuesten KI-Methoden zur Verfügung zu stellen. PrototAlp wird als Programmierumgebung konzipiert, in der Anwender*innen ihre eigene Software programmieren können. Der große Vorteil zu bestehenden Lösungen ist der, dass PrototAlp die wichtigsten und gängigsten Software-Bibliotheken bereits vorinstalliert enthält. Es muss nichts selbst installiert werden, Anwender*innen können sofort loslegen und die Eintrittsschwelle sinkt dadurch.

Außerdem wird PrototAlp als „Platform-as-a-Service“ entwickelt. Die Anwender*innen greifen online darauf zu und die Experimente und Berechnungen laufen auf den PrototAlp-Systemen. Dadurch sinken die Hardware-Anforderungen auf der Nutzer*innenseite. PrototAlp stellt auch leistungsfähige Grafikkarten zur Verfügung, die besonders gut geeignet sind für Anwendungen im Bereich KI, Machine-Learning, und Data-Science.

Im Folgenden beschreiben wir die Aktivitäten im ersten Berichtszeitraum. Wir beschreiben die Arbeiten zu Beginn des Projekts, die Durchführung von Use-Case-Evaluierungen, die Einrichtung der technischen Infrastruktur, sowie das Software-Setup.

2 Status der Arbeitspakete

2.1 Arbeitspaket 1 - Detailplanung und Formales am Projektstart

Kurzbeschreibung der Haupttätigkeiten. Die Haupttätigkeiten in diesem Arbeitspaket bestanden darin, organisatorische Arbeiten am Beginn des Projekts durchzuführen und die weitere Vorgehensweise ordnungsgemäß zu planen.

Erkenntnisse zur Vorgangsweise. Es wurden die formellen Aufgaben zum Projektstart durchgeführt (z.B. Abruf der ersten Förderrate). Weiters wurde die Webseite des Projekts auf der Plattform von netidee aufgesetzt sowie erste Ideen für die Blog-Post-Gestaltung gesammelt. Weiters wurde ein internes Projekt-Kickoff geplant, das am 04.03.2022 stattfand.

Kurzbeschreibung der erreichten Ergebnisse. Die Einrichtung der Webseite und der Abruf der ersten Förderrate wurden durchgeführt. Es wurde ein vorläufiger Blog-Post-Plan erstellt.

Besondere Erfolge/ Probleme. Der Erfolg entsprach den Erwartungen, es gab keine Probleme.

Gab es große Abweichungen zum Plan? Warum? Die erwarteten Ergebnisse wurden erreicht, es gab keine Abweichungen zum Plan.

2.2 Arbeitspaket 2 - Use-case-Evaluierung

Kurzbeschreibung der Haupttätigkeiten. Dieses Arbeitspaket hatte das Ziel, die Use-Cases und Bedarfe von späteren Nutzer*innen genauer kennenzulernen.

Erkenntnisse zur Vorgangsweise. Wir konzipierten eine Umfrage, die uns genauere Einblicke in das Nutzungsverhalten späterer Anwender*innen liefern sollte. Diese Umfrage umfasste Fragen zur Anwendungsdomaene (z.B.: Artificial Intelligence, Computer-Human Interfaces, NLP, ...), die Arten von Datensätzen, mit denen unsere Respondent*innen schon gearbeitet haben, ihre Hardware-Bedürfnisse, sowie bekannte und bevorzugte Software-Bibliotheken.

Kurzbeschreibung der erreichten Ergebnisse. Der Fragebogen wurde mit Google-Forms erstellt und innerhalb unseres eigenen Netzwerks von Data-Scientists verschickt. Es gab 7 Respondenten, die den Fragebogen ausgefüllt haben. Diese Zahl ist zugegebenermaßen gering, liefert aber dennoch wichtige Erkenntnisse für die weitere Gestaltung von PrototAlp, die zusammen mit unseren eigenen Erkenntnissen und einer Desk-Research ergänzt werden.

Besondere Erfolge/ Probleme. Das Umfrageergebnis hat bestätigt, dass Anwender*innen von der Verfügbarkeit von leistungsfähigen Grafikkarten profitieren können. Abbildung 1 zeigt das Resultat auf die Frage, ob eine starke GPU einen positiven Effekt auf die Arbeit der Respondenten hätte, die mit großer Mehrheit mit „Ja“ beantwortet wurde.

9) If you had access to a platform with a powerful GPU and preloaded libraries and datasets, do you think it would help you advance projects faster?

7 responses

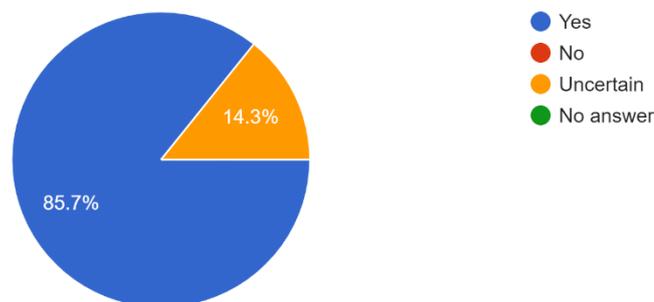


Abbildung 1 Laut Umfrage stellen GPUs eine signifikante Verbesserung bei der Durchführung von Data-Science-Experimenten dar.

In diesem Arbeitspaket gab es keine Abweichung vom Plan.

2.3 Arbeitspaket 3 - Setup und Management der Infrastruktur

Kurzbeschreibung der Haupttätigkeiten. In diesem Arbeitspaket wurde die technische Infrastruktur von PrototAlp aufgesetzt.

Erkenntnisse zur Vorgangsweise. Das Arbeitspaket umfasste Tätigkeiten wie die Auswahl einer für die Skalierbarkeit geeigneten Grafikkarte, das Einrichten des Betriebssystems auf dem Host-

Rechner, und die Installation notwendiger Software-Pakete für das Management der einzelnen Komponenten von PrototAlp.

Kurzbeschreibung der erreichten Ergebnisse. Ein Server wurde als Host fuer PrototAlp eingerichtet und die notwendigen Hard- und Software-Infrastruktur wurde aufgesetzt. Es wurde Docker als geeignete Lösung für die Orchestrierung von Software-Komponenten ausgewählt.

Besondere Erfolge/ Probleme. Der besondere Erfolg bestand darin, dass Dockerhub effizient als CI/CD-Umgebung genutzt werden konnte. Einerseits wurde dadurch uns Docker-Image öffentlich zugänglich gemacht, andererseits konnten wir so eine Lösung schaffen, mit der Nachinstallationen von Software einfach durchgeführt werden können.

Gab es große Abweichungen zum Plan? Warum? In diesem Arbeitspaket gab es eine Abweichung vom Plan: im Antrag wurde die Anschaffung einer leistungsfähigen Grafikkarte budgetiert. Diese Anschaffung wurde jedoch nicht benötigt, da ein Sponsor für diese Grafikkarte gefunden wurde. Das dafür vorgesehene Budget wurde in Personalkosten umgewandelt und so die Entwicklung von PrototAlp vorangetrieben.

2.4 Arbeitspaket 4 - Software-Setup

Kurzbeschreibung der Haupttätigkeiten. Das Software-Setup umfasste die Auswahl einer geeigneten Open-Source-Plattform, die für den Einsatz als gehostete Plattform verwendet werden kann, sowie die Auswahl und das Deployment von Software-Libraries, die für Anwendungen im Bereich Data-Science, Machine-Learning und künstliche Intelligenz geeignet sind.

Erkenntnisse zur Vorgangsweise. Mit Hilfe von Desk-Research wurden geeignete Plattformen identifiziert und der engere Kandidat dann in einem prototypischen Deployment evaluiert. Die weiteren Software-Bibliotheken wurden durch eine Online-Umfrage sowie ebenfalls Desk-Research identifiziert (siehe Arbeitspaket 2).

Kurzbeschreibung der erreichten Ergebnisse. Die Open-Source-Plattform „JupyterHub“ wurde als Backbone für PrototAlp ausgewählt. JupyterHub ist eine flexible und leicht erweiterbare Lösung, die unseren Anforderungen erfüllte. Weiters wurde ein Docker-Image erstellt, in dem die Software-Pakete für die Anwender*innen hinterlegt wurden. Dadurch ist ein leichtes Deployment (Installation) dieser Pakete in JupyterHub möglich, diese sind sofort für die Anwender*innen innerhalb der PrototAlp-Plattform verfügbar.

Besondere Erfolge/ Probleme. JupyterHub übertraf unsere Erwartungen als Grundlage für PrototAlp, weil es von Haus aus ein User-Management, Authorization, und die Möglichkeit für die Hochskalierung durch eine integrierte Container-Lösung beinhaltet.

Gab es große Abweichungen zum Plan? Warum? Es gab keine Abweichungen vom Plan, außer jener, dass statt JupyterServer die Lösung JupyterHub verwendet wurde, die jedoch gegenüber Ersterer deutliche Vorteile aufweist.

3 Umsetzung Förderauflagen

Es gab keine speziellen Förderauflagen.

4 Zusammenfassung Planaktualisierung

Alle Anpassungen des Plan-Excels kurz zusammengefasst

Im Excel wurde Stundendokumentation und die Arbeitspakete an den aktuellen Projektverlauf angepasst. Weiters wurden die Hardware-Anschaffungskosten entfernt. Diese sind stattdessen in Personalkosten umgewandelt worden.

5 Öffentlichkeitsarbeit/ Vernetzung

Beschreibung der bereits erfolgten Öffentlichkeitsarbeit oder Vernetzung, bzw. Beschreibung des Plans künftiger Aktivitäten

Das Projekt wurde im Berichtszeitraum mit Hilfe von Blogposts, deren Verlinkung auf der firmeneigenen News-Seite, sowie einer firmeninternen Road-Show, der Jahresklausur, mit einem Stand beworben. Durch diesen Stand wurde auch der Sponsor für die Grafikkarte gefunden, der in der Beschreibung von Arbeitspaket 3 erwähnt wurde.

6 Eigene Projektwebsite

Es gibt noch keine eigene Projektwebseite.