



netidee

PROJEKTE

CrOSSD2 - Extending the Critical Open-Source Software Database

Zwischenbericht | Call 19 | Projekt ID 7399

Lizenz: CC BY-SA

Inhalt

1	Einleitung.....	3
2	Status der Arbeitspakete	3
2.1	Arbeitspaket 1 - Detailplanung und Formales am Projektstart.....	3
2.2	Arbeitspaket 2 - Projektmanagement	4
2.3	Arbeitspaket 3 - Weiterentwicklung CrOSSD-Plattform und Metriken	4
2.4	Arbeitspaket 4 - Community-Arbeit und qualitative Erhebungen	5
2.5	Arbeitspaket 5 - Dokumentation und Formales am Projektende	6
3	Umsetzung Förderauflagen	6
4	Zusammenfassung Planaktualisierung.....	6
5	Öffentlichkeitsarbeit/ Vernetzung.....	6
6	Eigene Projektwebsite.....	7

1 Einleitung

Im Projekt CrOSSD2 bauen wir auf die Ergebnisse von CrOSSD1 auf und entwickeln diese gezielt weiter. Während im ersten Projekt der Aufbau einer Datenbank im Vordergrund stand, die es erlaubt, kritische Open-Source-Softwareprojekte (OSS) zu identifizieren und deren Gesundheitszustand anhand quantitativer Indikatoren (z. B. Anzahl aktiver Entwickler:innen, Commits oder Releases) zu bewerten, setzten wir in CrOSSD2 bewusst neue Schwerpunkte:

- Einführung neuer Metriken auf Basis von NLP-/LLM-Techniken, um qualitative Dimensionen wie Verständlichkeit, Kollaborationskultur und Konfliktmanagement zu erfassen.
- Stärkere Community-Arbeit durch Interviews und Workshops, um praxisnahe Perspektiven zu erfassen und so eine tiefergehende Bewertung der Projektgesundheit zu ermöglichen.

Die bisherige Arbeit in CrOSSD1 hat gezeigt, dass rein quantitative Kennzahlen nicht ausreichen, um ein verlässliches Bild der Projektgesundheit zu gewinnen. Entscheidend für die Nachhaltigkeit und Stabilität von OSS-Projekten sind vor allem die Qualität der Kommunikation, Zusammenarbeit und Governance-Strukturen innerhalb der Community. Hier wollen wir mit CrOSSD2 ansetzen: Metriken auf Basis von NLP-/LLM-Methoden sollen qualitative Dimensionen wie Verständlichkeit, Inklusivität oder Konfliktlösung besser erfassen. Ergänzend dazu werden durch Interviews und Community-Workshops praxisnahe Perspektiven der Beteiligten integriert, die wir in CrOSSD1 noch nicht berücksichtigen konnten.

Auf dieser Basis richtete sich die bisherige Arbeit in CrOSSD2 auf die technische Weiterentwicklung der CrOSSD-Plattform und ihrer Metriken (AP3) sowie auf die systematische Community-Arbeit (AP4).

2 Status der Arbeitspakete

2.1 Arbeitspaket 1 - Detailplanung und Formales am Projektstart

Arbeitspaket 1 wurde bereits mit April 2025 erfolgreich abgeschlossen. Die folgenden Arbeiten wurden in diesem initialen Arbeitspaket durchgeführt:

- der Vertrag wurde unterschrieben,
- der Detailprojektplan wurde erstellt und abgenommen,
- eine detaillierte Liste der geplanten Projektergebnisse mit Lizenz und Ort der öffentlichen Bereitstellung wurde erstellt und abgenommen,
- die Projekt-Website ging in Betrieb und ein erster Blogeintrag wurde erstellt, und die erste Förderrate wurde beantragt.

2.2 Arbeitspaket 2 - Projektmanagement

Das Projektmanagement verlief bislang erwartungsgemäß und ohne größere Hindernisse, sodass hier keine Umplanungen oder Ähnliches notwendig waren.

Im Blog unter <https://www.netidee.at/crossd2> haben wir plangemäß bislang drei Beiträge veröffentlicht:

- Im Februar 2025 haben wir die Ziele unseres Folgeprojekts vorgestellt, sowohl bzgl. der geplanten Gesundheitsanalysen durch KI-Metriken als auch bzgl. der geplanten stärkeren Community-Einbindung: <https://www.netidee.at/crossd2/wie-geht-es-weiter-crossd2>
- Im April 2025 haben wir eine Posterpräsentation auf der Social Artificial Intelligence Night an der FH St. Pölten vorgestellt: <https://www.netidee.at/crossd2/crossd2-auf-der-saint-2025>
- Im August 2025 schließlich haben wir unsere Interviewstudie zur Arbeit mit Open-Source-Software vorgestellt, sowie zur Mitwirkung bei dieser aufgerufen: <https://www.netidee.at/crossd2/call-participation-oss-interview-study>

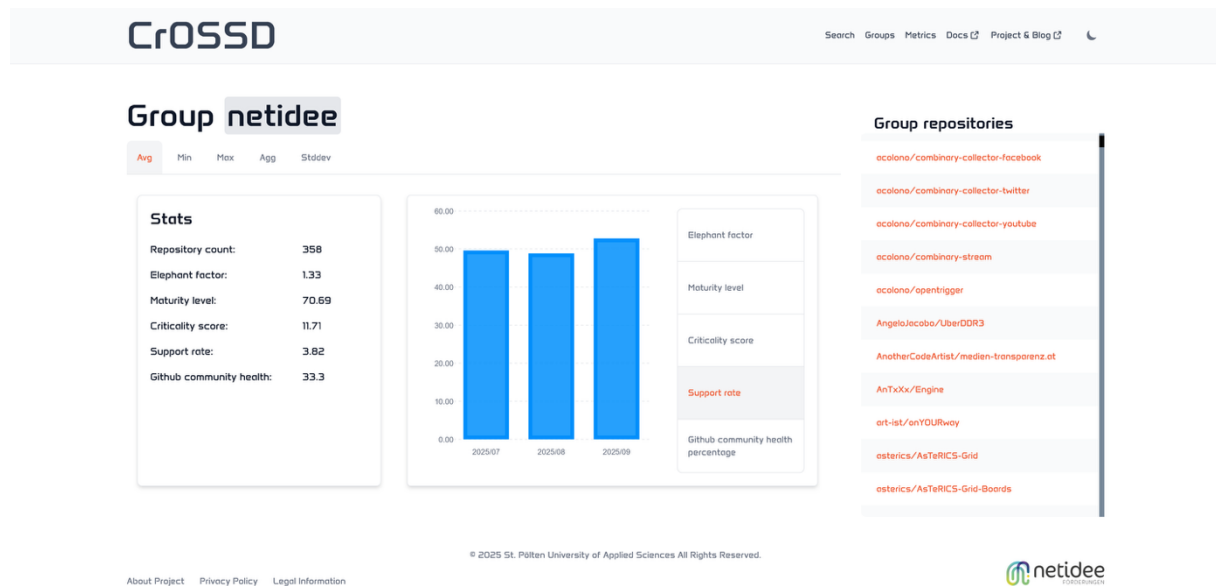
Auch Teil dieses Arbeitspaketes ist der fristgerechte Abschluss des Zwischenberichts.

2.3 Arbeitspaket 3 - Weiterentwicklung CrOSSD-Plattform und Metriken

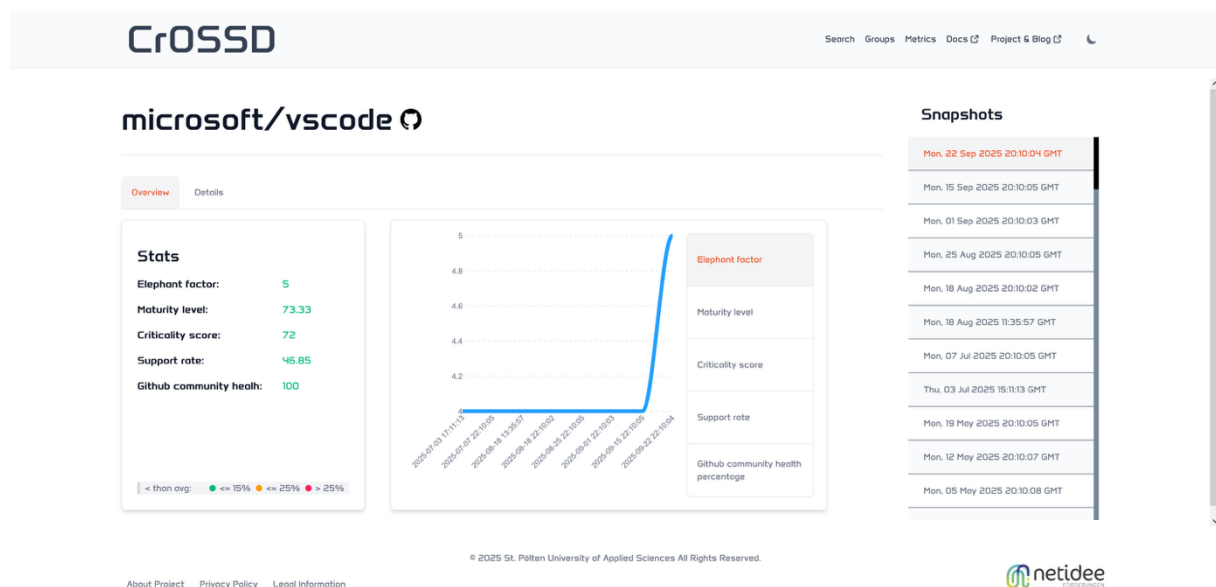
Um die Sicherheit der einzelnen Komponenten des CrOSSD Cluster gewährleisten zu können, wurden die benötigten Softwareabhängigkeiten auf aktuelle Versionen umgestellt und die nötigen Änderungen angewandt.

Der Crawler, welcher die benötigten Daten und Metadaten der Github Repositories sammelt, wurde grundlegend überarbeitet. Das Akquirieren der Daten wird soweit möglich über die Github GraphQL API erledigt, da hierbei weniger Request nötig sind als bei der Verwendung der entsprechenden Rest APIs. Dies reduziert die verursachte Last auf Seiten Githubs und beschleunigt die Scans im Allgemeinen.

Das Webinterface wurde auf die neue Version von Svelte und Sveltekit aktualisiert und das Ladeverhalten verbessert. Weiters wurde eine Gruppenansicht hinzugefügt. Diese ermöglicht zu sehen, welche Repositories zu einer Gruppe gehören, beispielsweise der Gruppe “netidee”, und welche Ausprägungen, wie etwa Durchschnitt, Minimum und Maximum, die wichtigsten Metriken in dieser Gruppe haben.



Die Projektseite der einzelnen Repositories wurde ebenfalls überarbeitet und bietet nun einen Überblick über die wichtigsten Metriken und deren Relation zu den Durchschnittswerten.



2.4 Arbeitspaket 4 - Community-Arbeit und qualitative Erhebungen

Im Laufe der letzten Monate wurden von uns über zwanzig Interview mit verschiedenen Stakeholder:innen geführt; hauptsächlich handelt es sich bei den Befragten um Softwareentwickler:innen und Programmierer:innen, aber auch einige in anderen Funktionen am OSS-Ökosystem teilnehmende Personen wurden befragt. Die Interview dauerten meist zwischen 30 und 60 Minuten, mit einigen wenigen Ausreißern um die 75–85 Minuten.

Der Interviewleitfaden fokussierte auf drei Kernpunkte: einerseits die benutzten Kriterien für Entscheidungen rund um die Verwendung und/oder Unterstützung von OSS-Paketen, andererseits die Verständnisse und Konzeptionen von Kritikalität, und zuletzt noch ein paar grundlegende Fragen rund um das Ziel, die erwartete Nützlichkeit und die grundlegende Zugangsweise von CrOSSD.

Einige wenige Interviews werden in den nächsten Wochen noch durchgeführt, insgesamt ist aber die theoretische Sättigung (trotz immer wieder hinzukommender einzelner interessanter Details oder Sichtweisen) bereits erreicht.

In den kommenden Monaten werden die Interview (unter Nutzung von Whisper, konkret noScribe) transkribiert und in der Folge induktiv analysiert, um die qualitativen Daten für die weitere Nutzung aufzubereiten.

2.5 Arbeitspaket 5 - Dokumentation und Formales am Projektende

(Dieses Arbeitspaket wurde noch nicht begonnen.)

3 Umsetzung Förderauflagen

(In der Fördervereinbarung sind keine Förderauflagen vorgesehen.)

4 Zusammenfassung Planaktualisierung

(Wir hatten bislang noch keinen Bedarf für Anpassungen im Projektplan.)

5 Öffentlichkeitsarbeit/ Vernetzung

Die bisherige Öffentlichkeitsarbeit und Vernetzung in CrOSSD2 umfasste folgende Aktivitäten:

- Veröffentlichung von 3 Blogposts im netidee-Blog;
- Präsentation des CrOSSD2 Projekts sowie ersten Ergebnissen auf der Social Artificial Intelligence Night 2025 an der FH St. Pölten.
- Social-Media-Aktivitäten auf den Accounts der Projektmitarbeiter sowie auf den Kanälen der FH St. Pölten, siehe z.B.: https://www.linkedin.com/posts/lukas-daniel-klausner-4aa484249_do-you-regularly-work-with-or-develop-open-source-activity-7363161315902664705-l2bZ?utm_source=share&utm_medium=member_desktop&rcm=ACoAABgJSWwBjL5-FHXpWKseEraZ5PotXyUMrGY

6 Eigene Projektwebsite

Die Projektwebsite, die bereits in CrOSSD 1 aufgebaut wurde, wird weiterverwendet:

<https://crossd.tech>

Die Live-Datenbank zur Berechnung und Darstellung von wichtigen/kritischen Open-Source-Software-Projekten und ihren aktuellen „Gesundheits“-Metriken wird kontinuierlich weiterentwickelt:

<https://health.crossd.tech/>