

Richtlinie zum „Software-Sprint“

Podcast Plugins

Schlussbericht

Zuwendungsempfänger:

Klaus Scheuermann

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS24S40 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Kurze Darstellung der Aufgabenstellung und Motivation

Was war Deine Motivation? Welches Problem wolltest Du mit Deinem Projekt lösen? Wie war die geplante Vorgehensweise zur Problemlösung (auch Angabe der wichtigsten Meilensteine)?

Aufgabenstellung und Motivation

Podcasts haben sich als wichtige alternative Medienform etabliert. Doch während große Produktionen über professionelle Klangbearbeitung verfügen, klingen viele unabhängige Podcasts schlechter, weil ihnen finanzielle und technische Ressourcen fehlen. Schlechter Sound beeinträchtigt aber die Verständlichkeit, verringert die Aufmerksamkeitsspanne und mindert die Reichweite.

Meine Motivation war es, dieses Problem zu lösen, indem ich einfach bedienbare Audio-Plugins entwickle, die speziell auf Sprachaufnahmen optimiert sind. Diese sollen unabhängig von Fachwissen und Erfahrung eine professionelle Klangqualität ermöglichen.

Mit **PodcastPlugins** möchte ich dazu beitragen, dass unabhängige Podcasts nicht mehr an schlechter Tonqualität scheitern. So wird ihre Reichweite erhöht, ihre Inhalte barrierefreier und ihre Wahrnehmung verbessert.

Geplante Vorgehensweise und Meilensteine

1. Anforderungen und Funktionsumfang definieren

- Welche klanglichen Probleme sollen gelöst werden?
- wie löse ich diese Probleme, wenn ich Sprachaufnahmen manuell bearbeite?
- Welche Funktionen sind essenziell?

2. Entwicklung der Audio-Algorithmen

- Dynamikbearbeitung, Frequenzkorrektur, Rauschreduktion
- Tests mit KI-gestützten Verfahren

3. Modulare Architektur und GUI-Design

- Programmierung und Test einzelner Module
- Gestaltung einer intuitiven Benutzeroberfläche

4. Integration und Plugin-Entwicklung

- Umsetzung als VST3-, CLAP- und LV2-Plugin
- Plattformübergreifende Tests auf Linux, macOS und Windows

5. Finalisierung und Veröffentlichung

- Erstellung eines Demo- und Tutorial-Videos
- Veröffentlichung als Open-Source-Projekt unter GPL-3 Lizenz auf github.com

Beitrag des Projektes zu den Zielen der Förderinitiative

„Software-Sprint“

Wer ist die Zielgruppe für Deine Lösung? Wie profitiert sie von den Ergebnissen Deines Projekts? Welche Bezüge gibt es zu den Themenfeldern und Zielen des Software Sprints?

Beitrag des Projekts zu den Zielen der Förderinitiative „Software-Sprint“

Zielgruppe und Nutzen

Die Hauptzielgruppe von **PodcastPlugins** sind unabhängige Podcast-Produzent:innen, die ihre Inhalte mit professioneller Klangqualität veröffentlichen möchten, aber nicht über teure Studio-Software oder tiefgehendes Audio-Wissen verfügen. Auch Journalist:innen, Bildungsmacher:innen und Aktivist:innen, die Podcasts als Medium nutzen, profitieren von der Lösung.

Durch einfach bedienbare Audio-Plugins ermöglicht **PodcastPlugins** eine deutlich bessere Sprachverständlichkeit, reduziert Ablenkungen durch schlechten Sound und erhöht somit die Reichweite und Wirkung nicht-kommerzieller Inhalte. Die Tools helfen

außerdem bei der Barrierefreiheit, indem sie für gut verständliche Sprachaufnahmen sorgen.

Bezug zu den Zielen des Software-Sprints

Das Projekt passt zur **Förderinitiative „Software-Sprint“**, da es:

- **Freie Software für gesellschaftlichen Mehrwert schafft: PodcastPlugins** wird als Open-Source-Software entwickelt, um Wissen und Technologie für alle zugänglich zu machen.

- **Technische Innovation mit einem klaren Praxisbezug kombiniert:** Durch die Nutzung der DSP-Sprache *Faust* und modularer Architektur werden komplexe Audio-Prozesse in eine intuitive Anwendung überführt.

- **Unabhängige Medien und digitale Souveränität stärkt:** Indem das Projekt den Zugang zu hochwertigem Sound demokratisiert, hilft es, mediale Vielfalt zu sichern und ermöglicht es mehr Menschen, sich mit Podcasts Gehör zu verschaffen.

Mit **PodcastPlugins** wird die Podcast-Landschaft diverser, inklusiver und professioneller – ohne finanzielle oder technische Hürden.

Ausführliche Darstellung der Ergebnisse

Welche konkreten Ergebnisse hast Du erzielt? Konnten alle Meilensteine erreicht werden? Welche zusätzlichen Erkenntnisse hast Du aus der Projektarbeit gewonnen, auch im Hinblick auf die Begleitung durch die Open Knowledge Foundation?

Die Ziele des Projekts wurden erreicht: Die Podcast-Plugins konnten zur Marktreife gebracht werden und stehen nun zur freien Nutzung in allen gängigen Formaten und plattformübergreifend bereit.

Im Förderzeitraum wurde zuerst an einem proof-of-concept gearbeitet, dessen Ziel es war, einen Algorithmus zu entwickeln, der automatisch und unabhängig vom eingehenden Audiomaterial eine klangliche Verbesserung vornimmt, die auf der Analyse meiner manuellen tontechnischen Arbeit beruht.

Dazu wurde ein mehrstufiger Prozess gebaut, der aus einem Pre-Filter, einem Spectral-Ballancer, einem Leveler, einem Multiband-Kompressor/Expander und einem Limiter besteht.

Der Pre-Filter bereinigt das Eingangssignal um Frequenzen, die für die Wiedergabe von Sprache nicht notwendig sind (Lowcut bei 60Hz, Highcut bei 16kHz).

Besonderes Augenmerk lag auf dem Design des Spectral Ballancers, da dort jene klanglichen Entscheidungen getroffen werden, die in den Hauptaufgabenbereich eines Tonmeisters oder einer Tonmeisterin bei der Bearbeitung von Sprache fallen (EQing).

Nach einigen Ansätzen stellte sich heraus, dass es eine universale Zielkurve zu geben scheint, an die sich das Frequenzspektrum des Eingangssignals anpassen soll um den gewünschten Effekt, nämlich ein angenehmes ausgeglichenes Klangbild, zu erzielen.

Entgegen anfänglicher Vermutungen, funktioniert diese Zielkurve für Stimmen aller Art und Tonlage. (Die Zielkurve kann im Quellcode nachgelesen werden.)

Technisch funktioniert die Frequenzanpassung (das Spectral-Ballancing) folgendermaßen: das Eingangssignal wird in 20 Frequenzbänder unterteilt. Von der wahrgenommenen Lautstärke (LUSF) jedes Bandes wird Wert der wahrgenommenen Lautstärke des Eingangssignals (full-range) subtrahiert. Dadurch bekommt man eine um die Lautstärke normalisierte Frequenzkurve des Eingangssignals, die nun mit der Zielkurve verglichen wird. Unterschiede werden im Folgenden durch Lautstärkeangleichung der einzelnen Frequenzbänder eliminiert. Der letzte Schritt folgt speziell dafür definierten Zeitkonstanten und Multiplikatoren, die dafür sorgen, dass die Angleichung als natürlich wahrgenommen wird.

Das Leveler-Modul wurde erst aus einem anderen Open Source Projekt (master_me) übernommen, dann jedoch deutlich verbessert. Das 'smoothing' der Lautstärkeangleichung folgt nun einem Paper von Andrew Smith, der darin den Begriff des 'dynamic smoothing' prägte. Daraus resultieren schnelle Lautstärkeänderungen, wenn die Eingangslautstärke weit von der Ziellautstärke entfernt liegt und langsame Änderungen, wenn der Unterschied klein ist.

Der Multiband-Kompressor/Expander ist ein klassisches Instrument der Sprachbearbeitung. Dafür wurden wohlklingende Parameter erarbeitet.

Selbes ist für den Limiter zutreffend.

Das Projekt ist mittlerweile auf GitHub veröffentlicht und erhält positives Feedback aus der Podcast-Community. Ein Promo- und Demo-Video auf YouTube findet Beachtung, und ein Toot zum Release erzielte auf Mastodon überraschend viel Reichweite.

Während des Förderzeitraums habe ich intensiv geforscht und viel gelernt. Besonders spannend war die Feinabstimmung der Zeitkonstanten im DSP-Algorithmus, da oft minimale Unterschiede darüber entschieden, ob der Klang natürlich oder künstlich wirkte. Die klanglichen Ergebnisse von PodcastPlugins haben mich positiv überrascht – die automatische Bearbeitung von Sprachaufnahmen funktioniert erstaunlich gut und spart erheblich Zeit und Ressourcen im Vergleich zur manuellen Nachbearbeitung.

Die Begleitung durch die Open Knowledge Foundation war eine große Unterstützung bei der Strukturierung des Projekts. Die regelmäßigen Zwischengespräche mit dem Team des Prototype Fund halfen, den Fortschritt des Projekts zu verfolgen und einen guten Überblick zu behalten. Besonders hilfreich war ein Coaching zur Konzeption von User Testings - ein Thema mit dem ich bis dato wenig Erfahrung hatte - das wertvolle Einblicke in die Benutzerfreundlichkeit der Software ermöglichte. Eine Coachin von Superbloom leitete mich an, standardisierte Fragen für potentielle Tester:innen zu erarbeiten, die bei darauf folgenden Testings und Demos abgefragt wurden.

Zielgruppe, Nutzen und mögliche Weiterentwicklungen

Welcher Nutzen ergibt sich für die Zielgruppe aus den Ergebnissen Deines Projekts? Welche weiter-gehenden Effekte ergeben sich aus der Open-Source-Stellung der Ergebnisse? Gibt es Ideen für die Weiterentwicklung Deiner Lösung und Pläne zu deren Umsetzung?

Hat die Arbeit in dem Projekt Dich in Deiner persönlichen, fachlichen Weiterentwicklung unterstützt?

Nutzen für die Zielgruppe

Die Zielgruppe – insbesondere unabhängige Podcaster:innen, Journalist:innen, Bildungsmacher:innen und Aktivist:innen – profitiert erheblich von **PodcastPlugins**, da es ihnen ermöglicht, ihre Inhalte mit professioneller Klangqualität zu produzieren, ohne tiefgehendes Audio-Wissen oder teure Studio-Software zu benötigen. Die Plugins sorgen für bessere Sprachverständlichkeit, angenehmeren Sound und eine höhere Konzentrationsspanne der Hörer:innen, wodurch Podcasts zugänglicher und effektiver werden. Zudem spart die automatisierte Klangoptimierung Zeit und Ressourcen, die sonst für manuelle Nachbearbeitung erforderlich wären.

Effekte der Open-Source-Stellung

Durch die Veröffentlichung als Open-Source-Software ergeben sich folgende positive Effekte:

1. Freie Verfügbarkeit & digitale Souveränität

- Nutzer:innen können die Software kostenlos verwenden, ohne an kommerzielle Anbieter gebunden zu sein.
- Unabhängige Produzent:innen erhalten Werkzeuge, die sonst nur großen Studios/Publishers vorbehalten sind.

2. Weiterentwicklung & Community-Beiträge

- Entwickler:innen können den Code verbessern, eigene Anpassungen vornehmen und neue Funktionen beisteuern.

- Die Community kann Fehler melden, Optimierungsvorschläge machen und so zur kontinuierlichen Verbesserung beitragen.

3. Transparenz & Datenschutz

- Im Gegensatz zu Cloud-basierten Alternativen läuft **PodcastPlugins** vollständig lokal, was die Kontrolle über Daten und Privatsphäre stärkt.

4. Nachhaltigkeit & Verbreitung

- Durch die Open-Source-Lizenz kann die Software langfristig gepflegt und weiterentwickelt werden, unabhängig von einzelnen Entwickler:innen oder Firmen.

Insgesamt stärkt **PodcastPlugins** die Unabhängigkeit von Medienschaffenden, trägt zur Demokratisierung hochwertiger Audioproduktion bei und fördert die Vielfalt in der Podcast-Landschaft.

Weiterentwicklung der Software

Auch wenn der DSP-Algorithmus sehr gut funktioniert, möchte ich mich dennoch kontinuierlich für die Weiterentwicklung und Verbesserung einsetzen.

Zudem möchte ich versuchen, weitere Anwendungsfelder für den den Speech-Enhancement Algorithmus zu ergründen. Denkbar wären Anwendungen in der digitalen Audiokommunikation, im Mobilfunk, oder bei KI generierter Sprache.

Persönliche fachliche Weiterentwicklung

Für mich persönlich entstand durch die Entwicklung der PodcastPlugins ein großer Kompetenz-Gewinn. Ich konnte neue DSP Ansätze erforschen und tiefer in die Plugin-Entwicklung eintauchen.

Kurze Darstellung der Arbeiten, die zu keiner Lösung geführt haben

Gab es Arbeiten bzw. Lösungsansätze, die nicht weiter verfolgt wurden? Was waren die Hintergründe, und wie bist Du alternativ vorgegangen?

Anfangs hatte ich die Idee, auch machine-learning in den Plugins zu verwenden. Dies erwies sich aber schon bald als weitgehend nutzlos, da klassische Audio-Prozesse deutlich besser funktionierten, als Versuche mit offenen KI-Modellen. Zudem basiert die Plugin-Architektur auf einem real-time Konzept, in dem sich KI nur mit sehr hohem Rechenaufwand nutzen lässt. Dies hätte zur Folge gehabt, dass wieder nur privilegierte Nutzer:innen von PodcastPlugins profitiert hätten.

Kurze Angabe von Präsentationsmöglichkeiten für mögliche Nutzer

Wo können sich Interessenten detailliert über Deine Projektergebnisse informieren (z.B. Webseite, GitHub, Veröffentlichungen)?

PodcastPlugins auf Github:

<https://github.com/trummerschlunk/PodcastPlugins>

Promo-/Demo-Video auf Youtube:

<https://www.youtube.com/watch?v=mGiD6aPvcZA>

Release-Post auf Mastodon:

<https://chaos.social/deck/@trummerschlunk/114093855059219251>

Kurze Erläuterung zur Einhaltung der Arbeits- und Kostenplanung

Gab es im Projektverlauf Ereignisse, die eine Anpassung der Planung erforderlich machten – z.B. Mehr- oder Minderaufwand bei der Bearbeitung von Teilaufgaben?

Nein.

Kurze Darstellung von etwaigen Ergebnissen bei anderen Stellen

Gab es Entwicklungen anderer Personen oder Institutionen, die Einfluss auf Deine Arbeiten und die Zielsetzung hatten? Wenn ja, worin bestand dieser und wie bist Du damit umgegangen?

Nein.