osemData – Modus Datenanalyst für die openSenseMap

Schlussbericht

Zuwendungsempfänger:

Umut Tas

DLR Projektträger Berlin
Eingegangen am
11. März 2020
Eingangsnr.: 0556 NM

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS19S12 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Kurze Darstellung der Aufgabenstellung und Motivation

Die Motivation hinter dem Projekt war es die Citizen Science Daten, die seit Jahren auf der openSenseMap gesammelt werden zugänglicher zu machen. Jeder Benutzer soll die Möglichkeit haben so einen Mehrwert aus den Daten zu ziehen und Theorien wie unser Handeln die Umwelt beeinflusst zu überprüfen. Dies war bis jetzt Benutzern vorbehalten die die Daten weiterverarbeiten konnten. Zudem sollte ein Augenmerk darauf gelegt werden das es den Benutzern Spass macht sich die Daten anzugucken und Hypothesen und Theorien über Umwelteinflüsse zu überprüfen. Die Entwicklung der neuen Platform basiert auf Umfragen unter den Benutzern der openSenseMap.

Beitrag des Projektes zu den Zielen der Förderinitative „Software-Sprint“

Die Zielgruppe sind alle Benutzer der openSenseMap, die sich in drei Gruppen aufteilen lassen. Zum einen die Interessierten Bürger, die gerne mehr über Ihre Umwelt wissen möchten und Teil der Citizen Science Kampange sind. Zudem wird die openSenseMap viel in der Bildung genutzt. Schüler lernen wie sie IoT-Geräte benutzen und Daten im Internet teilen. Jetzt haben Sie zusätzlich ein Werkzeug um die gesammelten Daten besser zu verstehen. Die dritte Zielgruppe wären Wissenschaftler, denen wir es ermöglichen wesentlich einfacher anomalien oder interessante Umweltphänomene zu entdecken und genauer zu untersuchen um so wissenschaftliche Fragestellungen beantworten zu können.

Ausführliche Darstellung der Ergebnisse

Es wurden zwei konkrete Ergebnisse während des Prototypefundes erzielt.

Zum einen ein Aufsetzen der Serverinfrastruktur auf einem OpenStack System (gehostet bei der Universität Münster) welches eine langfristige und zuverlässige Bereitstellung der Daten garantiert. Alle Skripte (Terraform, Ansible, Prometheus, etc..) wurden als OpenSource bereit gestellt und ermöglichem jedem eine Kopie der Infrastruktur selbst zu hosten.

Zum anderen wurde eine Platform entwickelt die eine ansprechende Dartstellung der Daten ermöglicht. Hierzu wurden eine Anpassungen an der API und dem Datenmodell der openSenseMap implementiert. Der Großteil der Arbeit floss in eine Frontend-Applikation die auf dem Angular Framework basiert und die am dringlichsten gewünschten Funktionen der Benutzer implementiert, wie z.B. Vergleiche, Zeitliche Entwicklungen oder räumliche Zusammenhänge.

Zielgruppe, Nutzen und mögliche Weiterentwicklungen

Die Zielgruppen und Nutzen wurden bereits auf Seite 1 erläutert.

Da die Ergebnisse noch nicht produktionionsreif sind und momentan nur eine „eingeschränkte" Version online verfügbar ist wird die Weiterentwicklung des Prototypen zu einem fertigen Produkt die erste Weiterentwicklung sein. Anschliessend sind weitere Arbeiten geplant die nicht mehr in den Rahmen dieses Prototypefundes gepasst haben, vorallem die Vereinheitlichung der Daten und anschliessend ein System um die Daten zu Klassifizieren um es den Benutzern zu ermöglichen genauere Analysen zu treffen (z.B. benötigen einige Fragestellungen Daten mit höherer Genauigkeit als andere, wir möchten allen Benutzern die Möglichkeit geben die Daten ensprechend zu filtern). Die Grundlage um diese beiden Probleme anzugehen wird ein Sensor-Wiki sein, welches Phänomene, Einheiten und Sensoren semantisch definiert.

Zudem befinde ich mich im Gespräch mit anderen Platformen die Umweltdaten anbieten um die Ergebnisse dieser Förderung dort zu Nutzen (e.g. Luftdaten.info, Sensor.Community).

Für meine persönliche Entwicklung war die Arbeit an diesem Projekt sehr positiv. Da ich eher als Frontend-Entwickler gearbeitet habe in meiner Laufbahn, war die Entwicklung der Infrastruktur super interessant, anspruchsvoll und im endeffekt aber auch sehr Lehrreich. Zudem kam ich durch die Treffen bei der OKFN in Kontakt mit Personen die an ähnlichen Projekten arbeiten und vorallem auch ähnliche Ideologien mit mir Teilen, die ich vermutlich sonst niemals getroffen hätte.

Kurze Darstellung der Arbeiten, die zu keiner Lösung geführt haben

Es gab zwei grössere Ansätze die ich nicht weiter verfolgt hab während der Förderung (die für das Projekt aber weiterhin wichtig sein werden in der Zukunft).

Zum einen die Umstellung der Datenbank auf eine Time-Series Datenbank mit der es wesentlich effizienter wäre die Daten abzurufen. Diesen Ansatz habe ich nicht weiter verfolgt, weil es einfach zu viel Arbeit wäre und den Rahmen des Prototypefundes gesprengt hätte. Stattdessen habe ich versucht die Menge der Daten die für die Visualisierungen gebraucht wird zu reduzieren.

Ein zweites Arbeitspaket das nicht weiter verfolgt wurde waren 3D-Visualisierungen, die zwar sehr gut aussehen und vielen Benutzern Spass machen, aber zu wenig Mehrwert lieferten um den Aufwand zu rechtfertigen. Die reine Darstellung der Daten war mit einfacheren Methoden klarer und besser realisierbar für alle Endgeräte.

Kurze Angabe von Präsentationsmöglichkeiten für mögliche Nutzer

https://sensebox.de – Das Projekt aus dem die openSenseMap entstanden ist, hier findet man allgemeine Informationen zur CitizenScience Kampagne.

https://docs.opensensemap.org – Beschreibung der API

https://data.opensensemap.org – Prototyp des während der Förderung entwickelten Frontends

https://github.com/sensebox/openSenseMap-infrastructure – Die Infrastruktur bei Github

https://github.com/sensebox/openSenseMap-api – Die API bei Github

https://github.com/sensebox/openSenseMapX – Das Frontend bei Github

Kurze Erläuterung zur Einhaltung der Arbeits- und Kostenplanung

Es kam ein Arbeitspaket unerwartet dazu, dafür habe ich auf ein anderes verzichtet. Im Projektplan hatte ich von einer „Bereinigung der Daten“ und Anwendung von statistischen Methoden um diese besser nutzbar zu machen gesprochen. Dieses Arbeitspaket ist aufgrund der Infrastruktur aus dem Projektplan gefallen, weil die bestehende Infrastruktur so leider nicht weiterverwendent werden konnte mussten wir den Umzug zum OpenStack der Universität Münster implementieren. Ohne Server von denen man die Daten abrufen kann wären die Besten statistischen Methoden nichts Wert gewesen.

Kurze Darstellung von etwaigen Ergebnissen bei anderen Stellen

-