Richtlinie zum „Software-Sprint“

Lok_Finz_EU – Das lokal finanzierte Europa

Schlussbericht

Zuwendungsempfänger:

Bada und Bonis und Krich GbR

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS18S19 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)?

Für das Projekt "Das lokal finanzierte Europa" war unsere Hauptmotivaton, alle Förderungen zu sammeln, zu standardisieren und zu geolokalisieren, die die 28 EU-Staaten aus den Europäischen Strategie- und Investitionsfonds (ESIF) in der Periode 2007-2013 erhalten haben. Obwohl die Transaktionen selbst gemäß den Anforderungen der Ratsverordnung veröffentlicht wurden und Open Knowledge Germany erhebliche Anstrengungen unternommen hat, um die öffentlich verfügbaren Daten zu sammeln und zu standardisieren, enthält keine der Lösungen die Standorte, die zu den Transaktionen gehören. Das Hauptziel unseres Projekts bestand darin, jeder Transaktion den nächstgelegenen Standort zuzuweisen. Zu diesem Ergebnis führten die folgenden Meilensteine:

1. Alle öffentlich verfügbaren Subventionsgeschäfte für die 28 EU-Mitgliedstaaten (.pdfs / Online-Datenbanken / .csv-Dateien usw.) aufspüren.

2. Alle Transaktionen in ein einheitliches, standardisiertes Datenbankschema laden.

3. Jede Transaktion anhand der in der Quelle verfügbaren Daten dem nächstgelegenen Standort zuordnen. Die Nähe der Standorte könnte wie folgt sein: national, regional, kreisangehörig, kreisnah, kommunal.

4. Die Daten veröffentlichen und einen interaktivan Plan für das Publikum erstellen.

Beitrag des Projektes zu den Zielen der Förderinitative „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?

Die Hauptzielgruppe unseres Projekts sind die 511 Millionen Bürger der Mitgliedsstaaten der Europäischen Union. Wir sind der Ansicht, dass die Bereitstellung dieses Datensatzes selbst nicht ausreicht, um das Bewusstsein zu stärken und Interesse zu wecken: Für die Durchschnittsbürger sind 319,59 Milliarden Euro, die in der Periode 2007-2013 aus den ESIF verteilt worden sind, unverständlich, da sie nicht in der Lage sind, so große Beträge zu begreifen. Daher hielten wir es für notwendig, die erhaltenen Mittel vergleichbar zu machen und die detaillierteste statistische Einheit zu verwenden, die für den Vergleich zur Verfügung steht: die ~ 115.000 Gemeinden der EU. Auf diese Weise wurde das gesamte Programm in kleinere Teile zerlegt. Die Bürger können so die Finanzierung ihrer nächsten Umgebung mit der ihrer Nachbargemeinden, die eigenen Werte mit nationalen Durchschnittswerten oder sogar mit den von fremden Dörfern oder Städten vergleichen. Dies hilft unserer Meinung nach der Anerkennung der Vielfalt der Europäischen Union in allen Ländern.

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?

Wir konnten unsere ersten drei Meilensteine vollständig und unseren vierten Meilenstein teilweise erreichen - dieser letzte ist ein work in progress. Wir sammelten Transaktionen von über 1 Million Begünstigten und wir standardisierten die Daten gemäß unserem etablierten Datenschema. Wir haben die Daten unter den ~ 115.000 Gemeindestandorten verteilt. Wir haben eine Pro-Kopf-Kennzahl erstellt, bei der wir die jeweilige Einwohnerzahl der Gemeinden in der EU 28 verwendet und den verteilten Datensatz auf den statistischen Ebenen LAU1, NUTS3, NUTS2, NUTS1 und NUTS0 zusammengefasst haben. Wir haben auch eine Geometriedatei gesammelt, die die Grenzen der Gemeinden darstellt. Diese Geometriedatei wurde in die oben genannten Grenzen der statistischen Einheiten konvertiert. Die Daten wurden mit ihrer Hilfe auf verschiedenen Ebenen durch einen benutzerdefinierten Kartenserver visualisiert. Der Datensatz wurde in eine Datenbank geladen und für die Abfrage vorbereitet. Es wurde ein responsives Website-Design erstellt und ein Frontend-Framework für die Veröffentlichung vorbereitet. Das Interaktionsdesign und die Live-Version der Anwendung wurde noch nicht veröffentlicht.

Für die Verarbeitung der originalen Quelldateien und Datensätze haben wir eine selbstbeschreibende Datapipeline erstellt, das ist aus der Quelldatei zu öffnen.

Wir haben uns stark auf Open Knowledge Deutschlands Sammlung veröffentlichter Datenquellen verlassen. Leider konnten wir ihre Scrapings nicht verwenden, da gerade die Lokationen in ihren verarbeiteten Datensätzen fehlen.

Zielgruppe, Nutzen und mögliche Weiterentwicklungen

Welcher Nutzen ergibt sich für die Zielgruppe aus den Ergebnissen Deines Projekts? Welche weitergehenden 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?

Wir meinen, dass die Zielgruppe ein klares und objektives Bild davon haben kann, wie viel Geld ihre engste Umgebung erhält und wie es sich auf andere territoriale Einheiten ausdehnt - Gemeinden, Bezirke, Bezirke, Regionen und Nationen. Es ist besonders wichtig, ein klares und verständliches Factsheet über das größte EU-Förderprogramm zu erstellen, wenn man sich bedenkt, wie überpolitisiert dieses Thema in letzter Zeit unter den 28 EU -Mitglied-staaten ist. (Wer ist der eigentliche Nutznießer solcher Programme? Welches Land erhält mehr als es einzahlt? War das Kohäsionsprogramm wirklich erfolgreich oder hat es nur die lokale Korruption gefüttert?)

Die Codebasis und, was noch wichtiger ist, die Daten selbst als Open Source zu verstehen, ist unseres Erachtens entscheidend für die Initiierung und Förderung weiterer Forschungen in diesem Bereich. Der Subventionsdatensatz könnte mit anderen statistischen, quantitativen und qualitativen Datensätzen kombiniert werden, um den Erfolg der EU-Subventionsprogramme in sehr präzisen, bisher beispiellosen Einzelheiten zu messen und die Korruptionsforschung zu unterstützen. Diese Methodik könnte auch im Finanzierungszyklus 2014 - 2020 angewendet werden, wenn die Daten veröffentlicht werden.

Für uns persönlich waren die letzten 6 Monate eine enorme Lernkurve: Die Aufgabe selbst schien unvorstellbar groß zu sein, fast ungreifbar. Um unser Ziel zu erreichen war unentbehrlich, grossen Wert auf die sorgfältige Zeiteinteilung und die Priorität der Aufgaben zu legen.

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?

Ursprünglich wollten wir die verarbeiteten Daten jeden Landes in einer einzigen Datei darstellen und wir erstellten die Datenpipeline, um ein Output zu liefern. Für einige Länder wurden diese Dateien so groß, dass es für einen durchschnittlichen Benutzer unmöglich gewesen wäre, mit ihnen zu arbeiten. Daher haben wir die Daten relational zerlegt - durch Fremdschlüssel verknüpfbar, um Doppelungen zu vermeiden - und so konnten wir den gesamten Datensatz auf ein Drittel seiner ursprünglichen Größe komprimieren. Dies bedeutete leider, dass wir die Codebasis jeden Landes umgestalten mussten. Wir hätten diesen Mehraufwand mit einer sorfältigeren Planung vermeiden können.

Kurze Angabe von Präsentationsmöglichkeiten für mögliche Nutzer

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

Der Open Source Code ist verfügbar unter:

https://github.com/eurospent/eurospent

Da die Daten sehr groß sind, veröffentlichen wir sie derzeit in einem Dropbox-Ordner:

Der mapserver befindet sich noch in einer Entwicklungsumgebung, erreichbar über:

Website-URL: http://map.eurospent.com

Wir haben auch eine Website für Analyse und Präsentation für den Endbenutzer vorbereitet. Dies ist ebenfalls im Aufbau, wird aber verfügbar sein unter:

URL der Website: http://eurospent.com/

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?

Wir haben die Erhaltungskosten sehr unterschätzt und uns überwiegend auf die Entwicklung des Projekts konzentriert. Daher müssen wir noch herausfinden, wie wir einen so enormen Datensatz kosteneffektiver verarbeiten und aktualisieren können.

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?

An diesem Teil müssen wir noch arbeiten. Nach der Fertigstellung der Webapplikation möchten wir vor eine EU-Media Kampagne zu starten und dabei das breite Publikum durch mindestens 10 wichtige Presserorgane anzusprechen. Ausserdem möchten wir unser Projekt an kleinere (regionale/lokale) Zeitungen bzw. Organe vorstellen und unser Zielpublikum auch online via Artikel im Internet erreichen.