Richtlinie zum „Software-Sprint“

Statistik Erfahrbar

Schlussbericht

Zuwendungsempfänger:

Name des Zuwendungsempfängers

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen **01IS19S17** gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei der Autorin.

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)?

Die siebte Runde des Prototype Fund brachte uns, das Team hinter dem Projekt Datenguide (gegründet in der zweiten Runde des Prototype Fund 2017/2018), zurück zu unserer ursprünglichen Mission: Statistiken für alle! Wir wollen die amtliche Statistik erklären und nutzbar machen – auch für Menschen ohne besondere Vorkenntnisse. Unsere großen Vorbilder waren dabei „Explorable Explanations“: interaktive, ins Auge fallende Visualisierungen komplexer Zusammenhänge.

Ziel des Projekts Statistik erklärt war es, herauszufinden, ob und wie wir solche interaktive Darstellungs- und Vermittlungsformen einsetzen können, um Hintergründe amtlicher Statistik so erklären, dass mehr Menschen sich trauen, damit zu arbeiten. Denn die Lernkurve zur ersten Datenvisualisierung oder zum ersten Blogpost, basierend auf amtlichen Daten, ist relativ steil. Um mit amtlichen Statistiken zu arbeiten, braucht es Wissen über Datenformate, über regionale Einteilungen und natürlich über Statistik. Das Problem: NUTS-Level, amtliche Gemeindeschlüssel oder Tabellenformate wirken wahlweise Furcht einflößend oder einschläfernd. In der Bevölkerung gibt es ein fehlendes Verständnis für Grundlagen der Statistik. Die New York Times hat z.B. einen Data Literacy Kurs für ihren Newsroom entwickelt. Viele Menschen begegnen Zahlen und Statistiken mit Unbehagen und Misstrauen. Dies betrifft zudem Frauen und andere unterrepräsentierte Gruppen, die seltener Berufe ergreifen, in denen statistisches Wissen vermittelt wird.

Um Daten spannend zu machen, braucht es Darstellungsformen, die Spaß machen. Unsere Explorable Explanations machen Statistiken haptisch, visuell und spielerisch erfahrbar. Diese neuen Darstellungsformen stehen dann zur Übernahme für andere Themen zur Verfügung.

Die geplanten wichtigsten Ziele waren eine Reihe von Artikeln im Web, die in einer Kombination mit interaktiven Datenvisualisierungen und spielerischen Bestandteilen Grundbegriffe der Statistik erklären und erfahrbar machen. Die Artikel und interaktiven Grafiken sollten für Beispiele und Visualisierungen offene Daten verwenden, insbesondere Daten der Statistischen Ämter. Dafür verwenden wir die Datenschnittstelle von Datenguide (https://datengui.de – 2017 gefördert durch den Prototype Fund unter dem Arbeitstitel “Datenland”).
Zweiter Meilenstein war darüber hinaus die Bereitstellung der visuellen und/oder interaktiven Komponenten zur Übernahme für andere Projekte. Ein Werkzeugkasten aus open-source-Komponenten für die Entwicklung von interaktiven Erklärstücken wie im ersten Meilenstein beschrieben. Damit soll die weitere Entwicklung solcher Anwendungen vereinfacht werden. Der Werkzeugkasten baut auf bestehenden open-source-Lösungen auf, insbesondere auf React, D3 und MDX (Markdown).

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?

Das Projekt “Statistik erklärt” (gefördert unter dem Arbeitstitel “Statistik erfahrbar”) trägt zu den Zielen zweier Schwerpunkte der Förderinitiative bei: “Data Literacy” sowie “Civic Tech”.
Im Sinne der “Data Literacy” wird die Einstiegshürde für Nutzer:innen ohne Vorkenntnisse deutlich verringert. Dies kann Berührungsängste vor amtlichen Statistiken nehmen und die Hemmschwelle, diese zu benutzen, senken. Unterstützt wird das davon, dass der Arbeitsaufwand für eigene Datenrecherchen und -aufbereitung für die Nutzer:innen erheblich gesenkt wird. Über ‘Statistik erklärt’ finden Nutzer:innen, die nach Fachtermini wie ‘NUTS-Level’ oder ‘Amtlicher Gemeindeschlüssel’ gesucht haben, zum Projekt Datenguide und werden direkt zum ebenfalls in diesem ‘Software Sprint’ veröffentlichten Prototypen des Datenguide Datenportals geleitet. Beide Projekt ermöglichen für den Bereich “Civic Tech” Bürger:innen einen deutlich vereinfachten Zugang zu Daten, die zwar grundsätzlich der Öffentlichkeit zur Verfügung stehen aber aufgrund der Art ihrer Veröffentlichung kaum von Bürger:innen verwendet werden können.

Ausführliche Darstellung der Ergebnisse

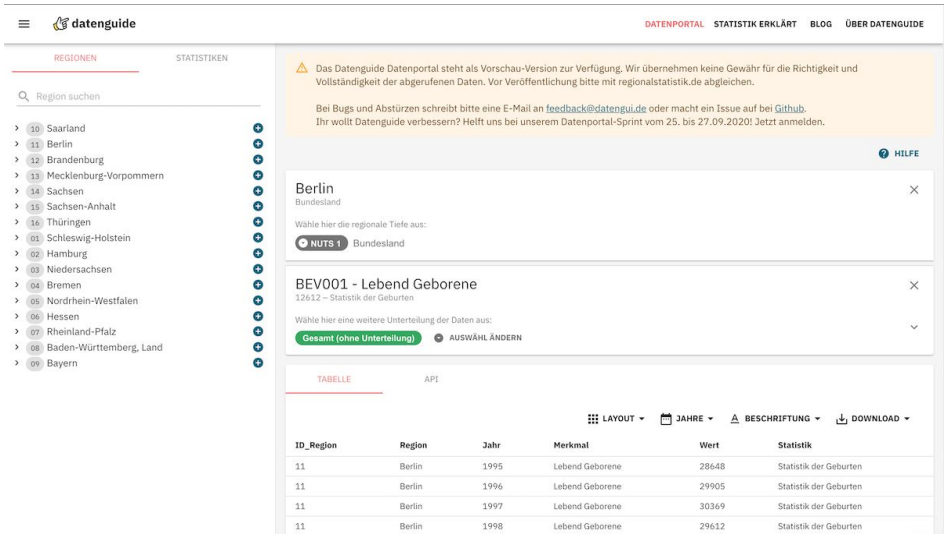
Welche konkreten Ergebnisse hast Du erzielt? Konnten alle Meilensteine erreicht werden? Welche zusätzlichen Erkenntnisse hast Du aus der Projektarbeit gewonnen, auch im Hinblick auf die Begleitung durch die Open Knowledge Foundation?

Für Datenguide sind die Ergebnisse dieser Prototype-Fund-Runde Meilensteine:

- Es gibt nun ein einfaches Tool, das zeigt, welche Daten für eine bestimmte Region verfügbar sind und mit dem auch Nutzer:innen ohne Vorkenntnisse auf die Daten der amtlichen Statistik zugreifen können.

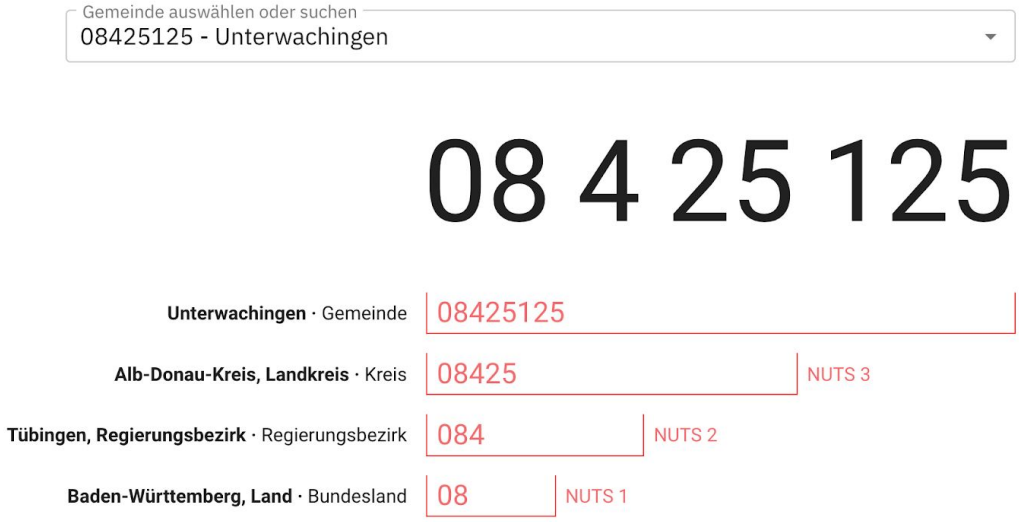
- Das Datenguide-Datenportal steht nun als Prototyp zur Verfügung. Darüber lassen sich die Daten aus regionalstatistik.de abrufen und in einem Format herunterladen, das sich einfach in anderen Visualisierungstools wie zum Beispiel Datawrapper weiterverarbeiten lässt.

- Die Schritte von der Datensuche über Tabellenformate bis zur ersten fertigen Visualisierung werden Nutzer*innen in einem Blogpost erklärt.



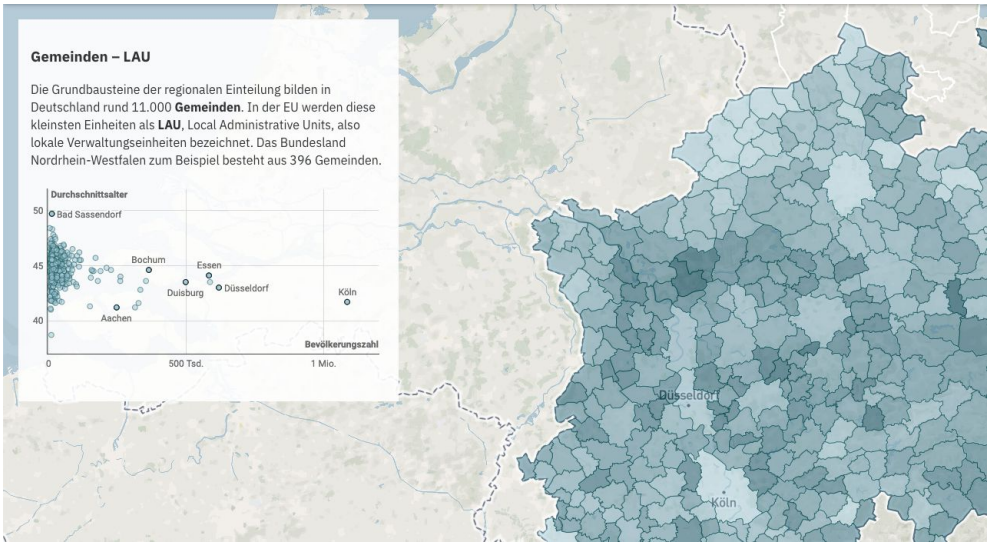
Datenguide ermöglicht es Nutzer:innen ohne Vorkenntnisse, amtliche Statistiken herunterzuladen.

- In einem weiteren Tool können Nutzer:innen den Namen einer Region eingeben, um den entsprechenden amtlichen Gemeindeschlüssel samt Erklärungen zurück zu bekommen.



Interaktive Erklärungen zum Amtlichen Gemeindeschlüssel (AGS)

- Ein Blogpost, erklärt interaktiv anhand von Karten, wie die europäische NUTS-Klassifikation für Deutschland funktioniert.



Interaktive Erklärungen zum NUTS-System

Alle diese Bausteine bieten Einstiegsmöglichkeiten in das Arbeiten mit amtlichen Statistiken und wir erhoffen uns, dass dadurch mehr Nutzer:innen zu Datenguide finden – auch über Suchmaschinen.

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?

Zur Zielgruppe: Bei Datenguide arbeiten wir an einer digitalen Plattform, die frei nutzbare Daten der statistischen Ämter durchsuchbar macht, klar und prägnant darstellt und es euch ermöglicht, diese Statistiken für eure eigene Analysen und Projekte weiterzuverwenden. Im Lauf der Zeit hat sich dabei der Fokus verschoben, hin zu einer Plattform für Expert:innen (wie Data Scientists und Datenjournalist:innen), die über Datenschnittstellen (APIs) auf Statistiken zugreifen. Mit dem neuen Projekt „Statistik erklärt“ laden wir auch Menschen ohne Vorkenntnisse einladen, amtliche Statistiken zu nutzen. Wir stellen dafür digitale Tools bereit, die helfen, regionale Statistiken zu verstehen und sinnvoll einzusetzen.

Ein weiteres Ergebnis des Projekts ist eine Sammlung von technischen Komponenten, mit denen interaktive Erklärstücke entwickelt werden können. Für die Umsetzung verwendeten wir JavaScript und React, sowie verschiedene spezialisierte Bibliotheken wie MapboxGL für Kartendarstellungen, MDX für das Einbetten von interaktiven Komponenten in Markdown-Dokumente und Scrollama für Scrollytelling-Interaktion. Sie können aber auch von allen weiterverwendet werden, der ähnliche Erklärstücke umsetzen möchten. Wir hoffen damit der deutschen Datenjournalisten-Community Impulse zu geben, sodass in Zukunft mehr ‘Explorable Explanations’ umgesetzt werden – gerne auch mit Daten aus der Regionalstatistik.

Im Rahmen der Projektarbeiten wurden wichtige fachliche Kenntnisse erlernt. So besteht nun bei den Projektteilnehmern ein umfassendes Wissen über amtliche Statistiken in Deutschland und deren Grundlagen, wie dem Amtlichen Gemeindeschlüssel oder der NUTS-Klassifikation. Zudem wurde musterhaft ein Prozess beschrieben, wie Nutzer:innen Daten per Datenguide abrufen und schnellstmöglich visualisieren können. Im Zuge der Programmierung der Explorable Explanations wurden neue Technologien erlernt und angewendet: Zum Beispiel Mapbox GL für die interaktive Karten in Kombination mit Scrollama für die Interaktivität.

Ebenfalls neue Erkenntnisse erlangten die Teilnehmer durch die wahrgenommenen Coachings im Bereich User Experience Design, die über die Open Knowledge Foundation angeboten wurden. Diese fachliche Weiterbildung kann nun auch im Arbeitsalltag angewendet werden.

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?

Zunächst war das Datenguide-Team sehr beeindruckt und inspiriert von den spielerischen Umsetzungen und technischen Möglichkeiten der „Explorable Explanations“. Das Start-Coaching durch den Prototype Fund unterstützte uns bei der Fokussierung auf die Nutzer:innen. Es ging darum herauszufinden, welches Wissen die Nutzer:innen brauchen, um amtliche Statistiken über Datenguide sinnvoll abrufen und nutzen zu können. Wir legten uns deshalb auf folgende Themen fest:

1. Wie funktioniert NUTS, die europäische Klassifizierung statistischer Regionen?
2. Wie ist der amtliche Gemeindeschlüssel aufgebaut und was sagt er uns über die Hierarchie der Regionen?
3. Welche Datensätze gibt es in der Regionaldatenbank und für welche Regionen sind sie verfügbar?
4. Wie bereitet man Daten aus der Regionalstatistik so auf, dass man sie als Grafiken darstellen kann?

So sinnvoll und nötig für die Arbeit mit Regionalstatistik diese Themen sind, ein wenig ernüchtert hat uns die inhaltliche Recherche auch. Denn zum Beispiel die Aufteilungen in Länder und Gemeinden wird zwar auf europäischer Ebene maximal verklausuliert NUTS genannt und ist ohne Frage nötiges Grundwissen, um sortiert und schnell auf die Daten zugreifen zu können – bei genauerem Hinblick ist sie aber nicht übermäßig komplex.
Dazu kam die Erkenntnis, dass es wirklich schwer ist, komplexe Interaktionen, die auf verschiedenen Ebenen (Text, Karte, Highlighting) miteinander verzahnt sein sollen, im Kopf oder auf Papier vorauszuplanen.

Wichtig war uns immer wieder die Rückbesinnung auf die Nutzer:innen. Ursprünglich geplant waren Long-Read-Artikel mit interaktiven Erklärungen. Wir merkten aber schnell, dass es zum Beispiel beim Amtlichen Gemeindeschlüssel (AGS) gar nicht so viel im Text zu sagen gab, sondern dass der Fokus eher auf einem personalisierten Tool liegen muss, in dem Nutzer:innen für ihre Gemeinde den AGS herausfinden können und die Erklärung dadurch mitgeliefert bekommen.

Für Oktober war die Verwendung der Komponenten durch das Datenteam des Westdeutschen Rundfunks und der Redaktion von ‘Die Story’ geplant. Dies wurde jedoch von Seiten der Redaktion nicht umgesetzt.

Kurze Angabe von Präsentationsmöglichkeiten für mögliche Nutzer

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

Webseite: https://datengui.de/
Statistik erklärt: https://datengui.de/statistik-erklaert
Datenportal: https://datengui.de/statistiken

Github:

- Datenguide-Website, inklusive der interaktiven Erklärstücke: https://github.com/datenguide/datenguide
- Unsere Sammlung von Komponenten für interaktive Erklärstücke: https://github.com/datenguide/explorables

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?

Ursprünglich geplant war die Umsetzung in gemeinsamen Sprints vor Ort in Berlin und Köln. Bedingt durch die Corona-Pandemie konnten diese Sprints nur via Videokonferenz stattfinden. Dies machte teilweise das gemeinsame Arbeiten schwerer und führte dazu, dass der Kostenplan zwar insgesamt eingehalten werden konnte, sich aber die Verteilung des Arbeitsaufwandes innerhalb des Teams deutlich verschob. Ursprünglich war die Drittelung des Aufwandes geplant, zuletzt verschob sich die Umsetzung deutlich in Richtung von Simon Jockers, unter anderem weil sich die technische Umsetzung der Scrollytelling-Karte mit MapboxGL für Kartendarstellungen und Scrollama für Scrollytelling-Interaktion deutlich mehr Entwicklungszeit benötigte, als wir geplant hatten.

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?

Während der Projektlaufzeit gab es ein Update der Genesis-Online-Datenbank des Statistischen Bundesamtes. Dort lassen sich nun auch so genannte ‘Flat’-Tabellen herunterladen. Dies entspricht einer Empfehlung, die das Projekt ‘Datenguide’ bei Gesprächen im statistischen Bundesamt geäußert hatte. Auch das Portal regionalstatistik.de soll zeitnah um diese Funktion ergänzt werden.
Allerdings bieten auch die ‘Flat’-Tabellen noch ausreichend Hürden, die das nahtlose Weiterverarbeiten der Daten erschweren, so dass das Datenguide Datenportal für die Nutzer trotzdem einen Mehrwert und verringerten Zeitaufwand bietet. Dies konnten wir auch in einem Blog-Beitrag aufgreifen.

Wir verfolgen die Weiterentwicklung der offiziellen APIs sehr interessiert und greifen diese auch gerne auf. Datenguide versteht sich nicht als Konkurrenzprodukt, sondern als Brücke und Initiative, um die Themen API, Datenformate, Zugänglichkeit voranzutreiben.