

Endbericht

Mehr!Medientransparenz (prj1184)

Anmerkung: Durch das iterative Vorgehen mittels SCRUM wurden im Zuge der APs 5-7 auch die Ergebnisse früherer APs immer wieder überarbeitet. Wesentliche Änderungen werden daher bei den entsprechenden Arbeitspaketen gelistet.

Arbeitspaket 1

Bei diesem Arbeitspaket ging es vorrangig um die Erhebung der Anforderungen aus Kundensicht. Zu diesem Zweck wurde zunächst ein Leitfaden für ein strukturiertes Interview entworfen und mit dessen Hilfe wichtige Stakeholder (z.B. Vertreter der KommAustria, Datenjournalisten, Betreiber ähnlicher Web-Seiten) interviewt. Die Ergebnisse wurden mit dem Einverständnis der Gesprächspartner aufgezeichnet und anschließend entsprechend der Struktur des Leitfadens dokumentiert. Aus diesen Protokollen wurden die wichtigsten Anforderungen extrahiert und als sogenannter Product-Backlog für die Umsetzung eingetaktet.

Die Ergebnisse der Interviews waren besser als erwartet und resultierten in rund 40 Anforderungen unterschiedlichen Reifegrads. In Planungssitzungen wurden diese Anforderungen entsprechend verfeinert und priorisiert.

Das Arbeitspaket erfolgreich zu 100% abgeschlossen.

Arbeitspaket 2

Bereits in der ursprünglichen Version der Web-Seite hat sich gezeigt, dass die Qualität der Daten eher mangelhaft ist, was sich gerade in unzähligen Schreibweisen von Medien zeigt. Dadurch wird eine korrekte Verdichtung der Daten erschwert. In Gesprächen mit der KommAustria hat sich gezeigt, dass hier bereits Maßnahmen zur Verbesserung gesetzt wurden. So gibt es einerseits eine Liste mit Mediennamen, die laufend aktualisiert und von der KommAustria zur Verfügung gestellt wird, andererseits steht meldepflichtigen Unternehmungen auch ein Web-Formular zur Verfügung, welches die eingegebenen Namen mit diesen Einträgen vergleicht. Dieses wird jedoch nur von kleineren Zahlern verwendet, da größere Organisationen diese Daten mittels Datei übermitteln und diese Meldungen nicht weiter geprüft werden.

Ziel dieses Arbeitspakets war daher die Schaffung von automatisierten Korrekturen fehlerhafter Daten. Dazu werden beispielsweise alle angegebenen Web-Adressen normalisiert, indem das Protokoll und das Kürzel „www“ entfernt werden und der Rest in Kleinbuchstaben umgewandelt wird (z.B: „<http://derstandard.at>“ bzw. „[www.derStandard.at](http://www.derstandard.at)“ wird einheitlich zu „derstandard.at“).

Es hat sich aber gezeigt, dass nicht alle Daten vollständig automatisiert korrigiert werden können. Bei einer Recherche sind wir auf das Tool „Open Refine“ gestoßen, das speziell für die Verbesserung der Datenqualität in Open Data Beständen entwickelt wurde. Mit Hilfe dieses

Werkzeugs werden die durch die RTR veröffentlichten Daten nachbearbeitet und korrigiert. Es hat sich damit auch gezeigt, dass trotz der Anstrengungen der RTR die Qualität noch immer verbesserungsfähig ist.

Die von uns korrigierten Daten werden über GitHub (<https://github.com/AnotherCodeArtist/medien-transparenz.at/tree/master/data>) zur Verfügung gestellt und mittlerweile auch von anderen Web-Sites (bspw. Dossier.at) verwendet.

Ursprünglich erfolgte die Datenübernahme relativ einfach: Die aktuelle CSV-Datei wurde von der RTR-Seite heruntergeladen und per File-Upload eingespielt und automatisch übernommen.

Mit der neuen Methode mussten die Daten heruntergeladen und mit den bestehenden Daten integriert werden. Anschließend wurden die Daten mittels OpenRefine korrigiert und von dort wieder exportiert. Dieser Datenexport kann dann hochgeladen werden.

Nachdem die RTR nun aber auch eine JSON-Schnittstelle anbietet, konnte zumindest der erste Schritt hier wieder automatisiert werden.

Fertigstellungsgrad: 100%

Arbeitspaket 3

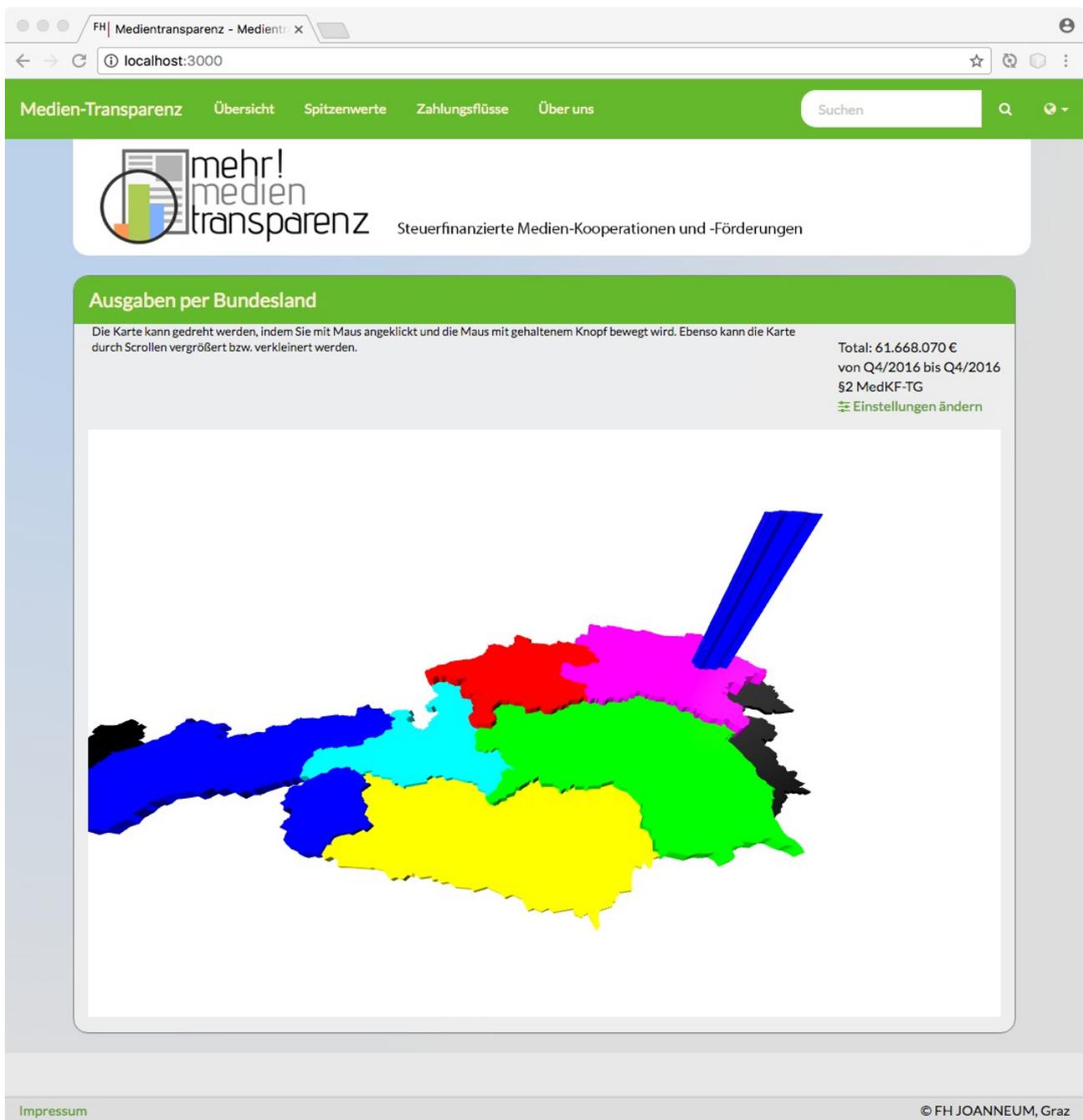


Abbildung 1: Verteilung der Ausgaben nach Bundesländern

Es ging darum, die Daten mit einem geographischen Kontext zu versehen. Dazu wurde ein Datensatz des Rechnungshofs ("Prüfobjekte") sowie Datensätze der Post (Beziehung zwischen Postleitzahlen und Bundesländern) verwendet (siehe Abbildung 1). Es hat sich aber gezeigt, dass eine reine Zurechnung zu Bundesländern irreführend ist, da sämtliche nationalen Organisationen (bspw. ÖBB) dem Bundesland Wien zugerechnet werden. Eine weitere Unterscheidung ist daher notwendig und wurde mittlerweile umgesetzt. Dazu werden die Einträge beim Importieren automatisch klassifiziert. Nach diesen „Klassen“ kann nun eingeschränkt werden (siehe Abbildung 2).

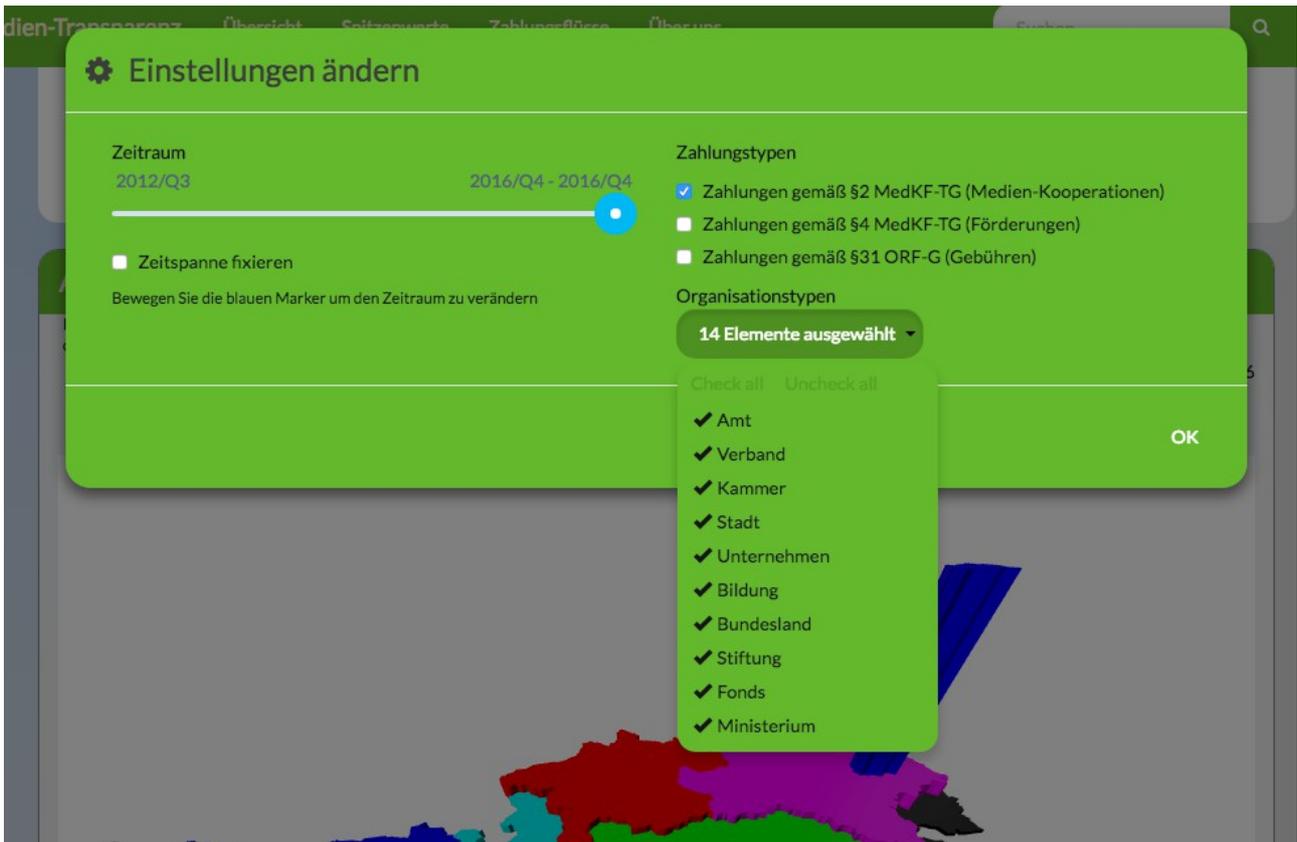


Abbildung 2: Klassifikation der Organisationen (Zahler)

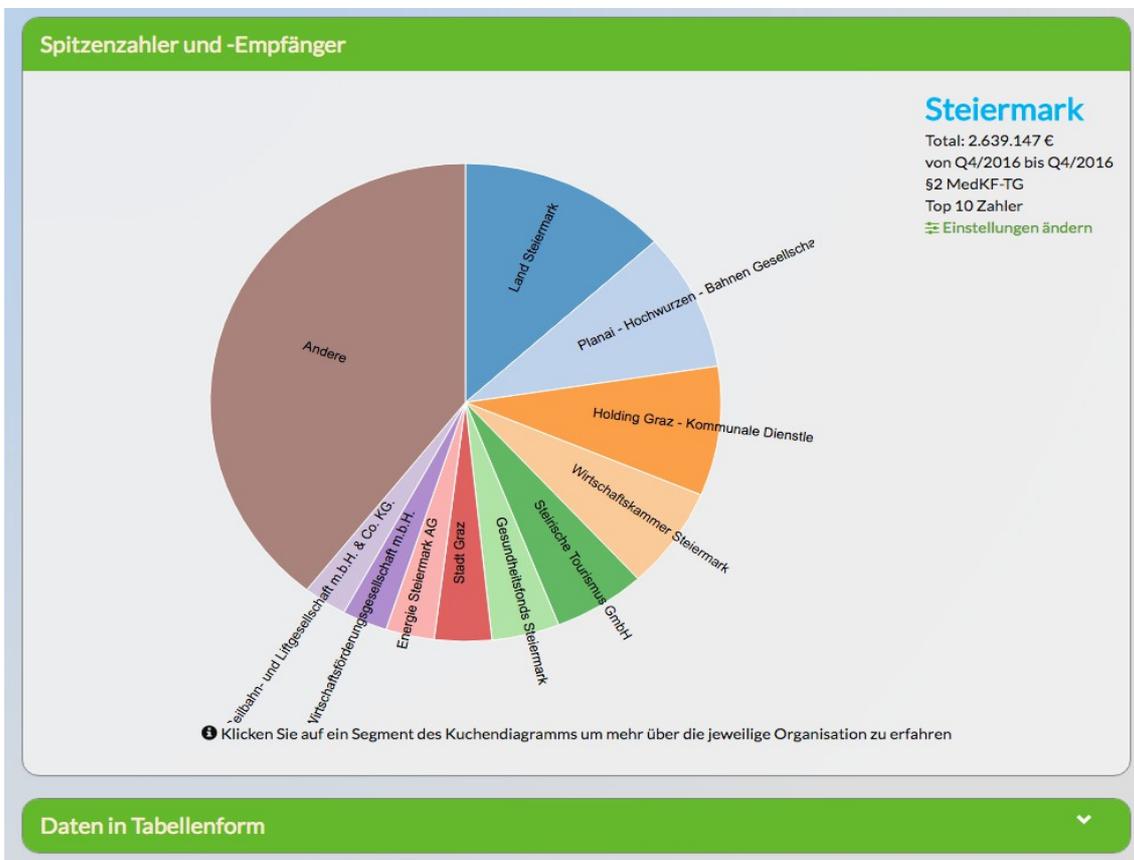


Abbildung 3: Top 10 Zahler aus Vorarlberg

Auch in anderen Ansichten kann nun nach Bundesländern gefiltert werden (siehe Abbildung 3).

Dieses Arbeitspaket ist nun ebenfalls zu 100% abgeschlossen.

Arbeitspaket 4

Hier wurde der initiale Product-Backlog erstellt. Es hat sich gezeigt, dass mit jedem neuen Zwischenergebnis wieder neue Ideen und Anforderungen entstanden sind, sodass der Backlog zahlreiche Änderungen durchlaufen hat.

Auch dieses Arbeitspaket ist zu 100% fertig

Arbeitspaket 5

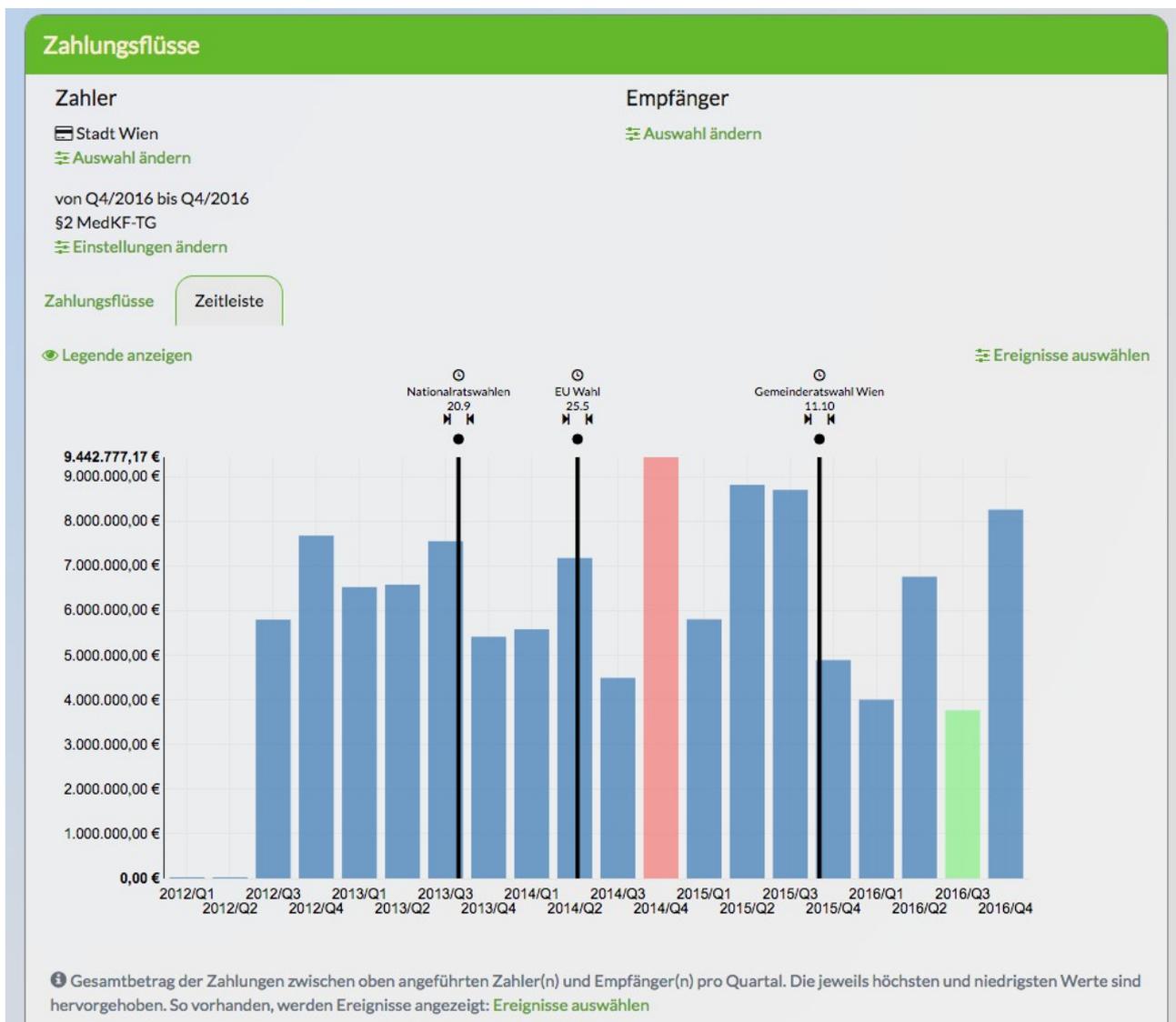


Abbildung 4: Zeitliche Entwicklung der Gesamtausgaben der Stadt Wien

Hier geht es um die eigentliche Umsetzung der identifizierten Anforderungen. Es wurden in Summe 15 Sprints durchgeführt, was auf intensive Arbeit und gerade das Finden neuer Ideen

zurückzuführen ist.

Es wurde eine Vielzahl neuer Features umgesetzt, welche eine wirklich detaillierte und einfache Analyse der Daten erlauben.

Eines der zentralen Features war hier die Abbildung des zeitlichen Verlaufs. Dieser Verlauf kann mit im System hinterlegten Events überlagert werden, um eventuell gegebene Zusammenhänge aufzuzeigen (vgl. Abbildung 4).

Hier kann einerseits der Verlauf der Gesamtausgaben, aber auch die Differenz der Ausgaben im Vergleich zum selben Quartal des Vorjahres dargestellt werden (Abbildung 5). Diese Variante wurde deshalb gewählt, da sich im Ausgabeverhalten Muster zeigen, die sich jährlich, quartalsweise wiederholen.

Alle Graphiken sind mit „Events“ überlagert, die einen möglichen Einfluss auf das Ausgabeverhalten haben könnten. Diese Events werden zentral verwaltet und können auf Basis unterschiedlicher Kriterien ein- oder ausgeblendet werden.

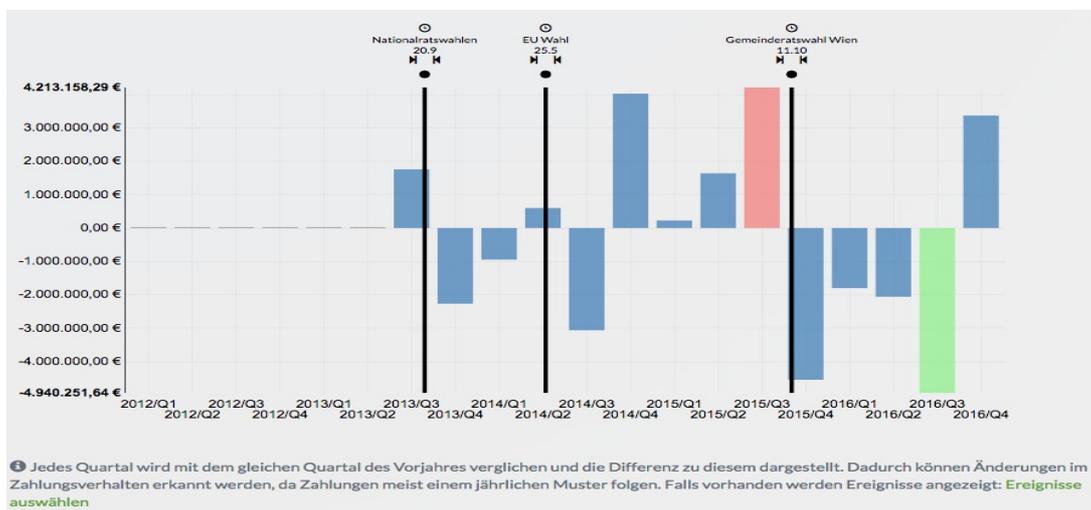


Abbildung 5: Differenz der Zahlungen im Vergleich zum jeweils gleichen Quartal des Vorjahres

Ein weiteres sehr zentrales Feature war die Möglichkeit zur Erstellung benutzerspezifischer Gruppen von Zahlern oder Empfängern (siehe Abbildung 6). Es gibt dabei zwei Typen von Gruppen. Systemweite, die nur von BenutzerInnen mit entsprechenden Rechten angelegt werden können, und lokale, die nur im Bereich der jeweiligen NutzerInnen sichtbar und gültig sind. Mit der Schaffung dieser Gruppen waren teilweise weitreichende Änderungen der Code-Basis verbunden, da diese in bestehende Ansichten integriert werden können.

Damit ist es nun möglich die Zahlungsflüsse zwischen Gruppen und anderen Organisationen zu analysieren, oder auch die Zahlungsflüsse zwischen Gruppen und deren wechselseitige Relevanz (siehe Abbildung 1).

Gruppentyp
 Gruppe für Organisationen Gruppe für Medien

Name
Styria

Eigentümer Organisation (Optional)
Geben Sie einen Namen ein

Zurücksetzen

Gruppenteilnehmer hinzufügen

Die Presse x Kleine Zeitung x Kleine Zeitung GmbH & Co KG x Kleine Zeitung Kärnten x Kleine Zeitung Steiermark x
Kleine Zeitung WOHN x PRESSE / TOP-1. BUCH x Presse x

Gruppe aktivieren
 Gruppe aktivieren Gruppe deaktivieren

Abbrechen Speichern

Abbildung 6: Dialog zum Anlegen und Ändern von Gruppen

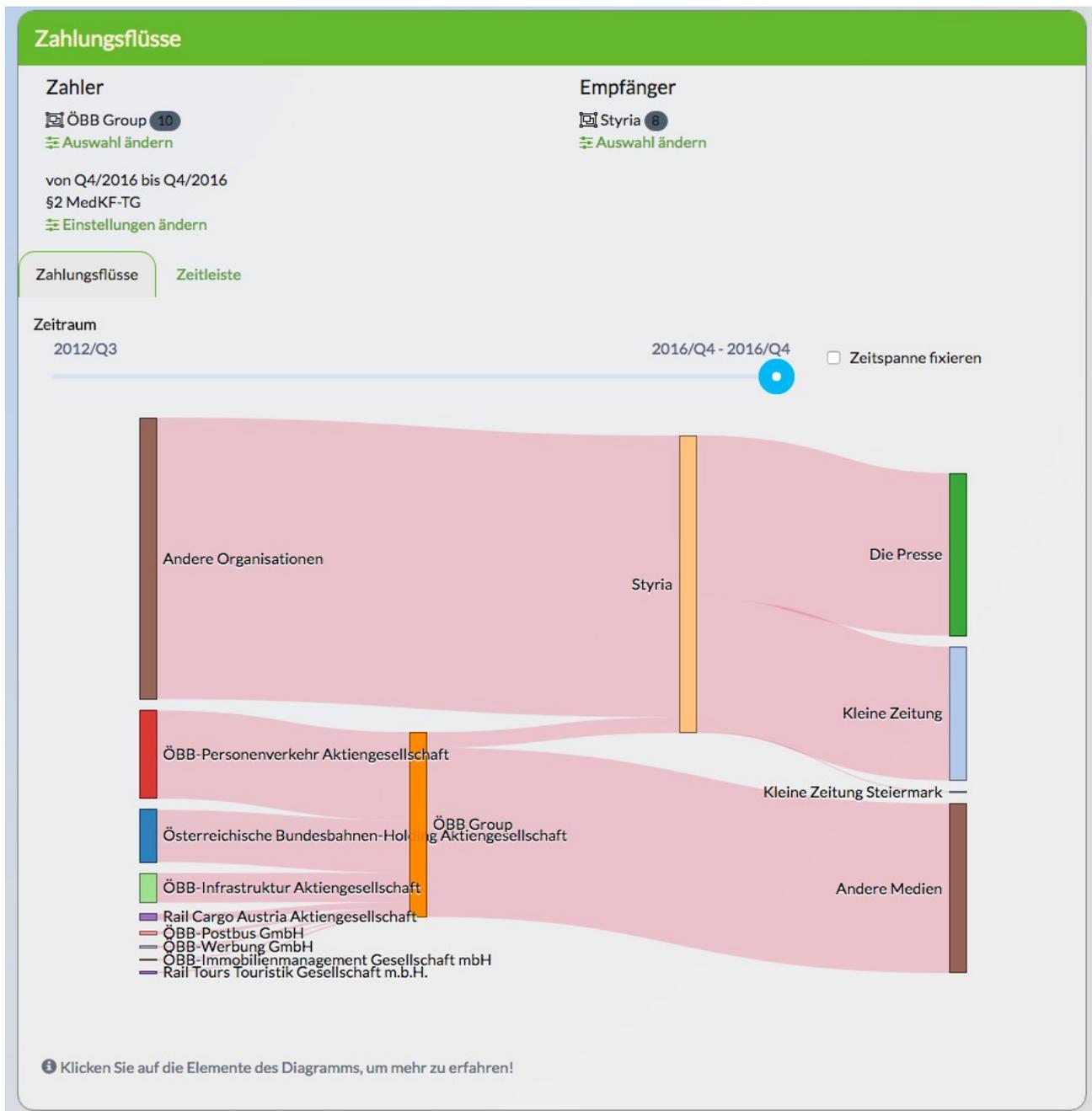


Abbildung 7: Zahlungsflüssen zwischen den Gruppen "ÖBB" und "Styria"

Das Arbeitspaket kann daher zu 90% als abgeschlossen angesehen werden.

Arbeitspakete 6+7

Diese Arbeitspakete wurden mehr oder weniger mit Arbeitspaket 5 integriert. Beim den letzten Sprints wurden jedoch keine Anforderungsänderungen mehr akzeptiert. Der Schwerpunkt der Arbeiten lag gegen Projektende auf der Gestaltung und Umsetzung eines neuen Designs. Dies führte letztendlich zu einer tiefgreifenden Änderung der Interaktion zwischen BenutzerInnen und Anwendung. Bis zur Umsetzung des neuen Designs, gab es auf jeder Seite eine Fülle von Einstellungsmöglichkeiten, die mit der Schaffung neuer Features ständig erweitert werden mussten. Diese wurden nun durch Links ersetzt (siehe Abbildung 8), die zu Einstellungsdialogen führen, die

wiederum an die gegebene Situation angepasst wurden (siehe Abbildung 9). Damit können BenutzerInnen die Möglichkeiten der Anwendung „entdecken“ und werden nicht vom Start weg mit einer Fülle von Optionen überfordert.



Abbildung 8: Links zu den Einstellungsdialogen



Abbildung 9: Einstellungsdialog

Arbeitspaket 8

Was die „Bewerbung“ bzw. Information über die modernisierte Plattform betrifft, wurde von unseren Journalisten ein Kommunikationskonzept entwickelt, das angepasste Schlüsselbotschaften und unterschiedliche Medien (einschließlich Kontaktadressen) sowie die Bewerbung der Projektergebnisse über diverse Homepages von Partnerorganisationen vorsieht.

Dieses Konzept wird in Kürze umgesetzt.

Projektergebnisse

Was die inhaltliche Komponente betrifft, so konnten wir alle unsere Ziele vollumfänglich erreichen. Es hat sich gezeigt, dass nach jedem Sprint neue Ideen zur Verbesserung bestehender Features entstanden sind und wir diese Umsetzung wollten. Es hat sich aber auch gezeigt, dass jedes neue Features auch teilweise massive Auswirkungen auf die bestehende Funktion hatte und dies zu umfangreichen Re-Factorings führte. Daher hat es auch immer wieder eine Weile gedauert, bis sich wieder ein stabiler Zustand eingeschwungen hat.

Sehr toll war für uns insbesondere die erste Projektphase, in der wir sehr viele tolle Vorschläge für die Verbesserungen und Erweiterungen erhalten haben.

Weniger toll war der Abgang eines wichtigen Entwicklers zu Amazon/Canada. Durch die Integration eines weiteren Wissenschaftlichen Mitarbeiters gegen Projektende konnten wir aber noch alles rechtzeitig fertigstellen.

Auch gibt noch einige Features in unserem Backlog, die wir gerne noch umsetzen möchten. Eventuell wird also wieder einen netidee-Antrag geben!

Erreichbarkeit

Die Website ist unter <https://www.medien-transparenz.at> erreichbar.

Der gesamte Source-Code (einschließlich der überarbeiteten Daten der RTR) kann via github bezogen werden:

<https://github.com/AnotherCodeArtist/medien-transparenz.at>

Die Software wurde unter der Apache2 License (<http://www.apache.org/licenses/LICENSE-2.0>) veröffentlicht.

Verwendete Open Source Software

Die technische Basis bietet das mean.io Framework, welches unter der MIT License veröffentlicht wurde.

MEAN ist dabei eine Abkürzung für

- MongoDB (GNU AGPL v3.0 siehe <https://www.mongodb.com/community/licensing>)
- Express (MIT License, siehe <https://github.com/expressjs/express>)
- AngularJS (MIT License, siehe <https://github.com/angular/angular.js>)
- Node JS (unterschiedliche Lizenzen für Teilkomponenten, siehe <https://github.com/nodejs/node/blob/master/LICENSE>)