Projektarbeit – Soziale Netzwerkanalyse

Datum: 7. November 2016

Team: Gabriel Brunner, Roger Matl, Jessica Odermatt, Gery Hirschfeld

# Thema

Was möchten Sie untersuchen? Was sind die Erwartungen?

Wir möchten Benutzer von GitHub und deren benutzten Technologien analysieren. Wir erwarten mehrere Cluster von User, die sich um ähnliche Technologien tummeln, da oft die gleichen Technologien miteinander kombiniert werden, z.b. HTML, CSS, JavaScript.

(optional) Zusätzlich möchten wir analysieren, welche der Benutzer einander „watchen“. Auch hier erwarten wir, dass die User in dem gleichen Cluster einander folgen, da die Interessen gleichartig sind.

# Fetcher

## Datenquelle

Beschreiben Sie, welche Datenquelle Sie verwenden. Erläutern Sie die verwendete Modellierung sowie auch, welche Attribute und Beziehungen abgebildet wurden.

GitHub API: <https://developer.github.com/v3/>

## Queries

Zeigen/ Beschreiben Sie die dafür notwendigen Queries.

## Einschränkungen

Wurden bereits irgendwelche Einschränkungen beim Sammeln der Daten definiert, um die Datenmenge gering zu halten? (z.B. nicht alle Personen, sondern nur Europäische Personen)

Filtern nach Schweizer Benutzer funktioniert nicht, da Location-Field im Profil ein textbasiertes Inputfeld ist. Müsste Auswahlfeld sein.

# Analyse

Schauen Sie im Abschnitt „Analyse“, welche Themen behandelt werden müssen. Erstellen Sie für jedes der drei Themen (Allgemeine Informationen, Filterung des Netz-werks, Analysen) einen neuen Unterabschnitt.

## Datenanalyse

Allgemein Informationen / Datenanalyse über das Netzwerk (Anzahl Knoten und Kanten), resp. welche Unsicherheiten vorhanden sind (falsche Informationen aufgrund fehlerhafter Quelle oder mangelnder Extraktions-Qualität)

## Filterung

Filterung des Netzwerks (ggf. auch schon beim Einsammeln) / Analyse verschiedener Views

## Analysen

### Graph-Metriken

### Zentralitäts-Masse und Prestige-Masse

### Clusters / Communities

# Fazit

Wurden Ihre Erwartungen erfüllt? Gab es spezielle Erkenntnisse?