

Do, 24.1.2013

Open Data-Workshop des VSt

# Open Data in Berlin: Stand und Erfahrungen



Block D 16.15 - 18.15

Hartmut Bömermann

# Inhalt

- Kurzer Abriss
- Umsetzungsprobleme
- Open Data-Beispiele
- Nutzen und Risiken

## Phase 1: Pilotprojekt Datenportal

- Nach **Vorstudie** und **Projektphase** mit einem externen DL:
  - **Pilotbetrieb** [daten.berlin.de](http://daten.berlin.de) ab 09/2011
  - 01/2012 **Bericht** „Berliner Open-Data-Strategie“
- Projektbeteiligte: SenWTF (FF), SenInnSport, AfS, Stadtportal BerlinOnline / Fraunhofer Fokus (DL)
- Ausgewählte **strategische Handlungsempfehlungen** des Berichts:
  - Politisches Bekenntnis zu offenen Daten
  - Gesamtverantwortlichen für offene Daten und Ansprechpartner für die Daten bestellen
  - Identifizierung der Daten und Workflow entwickeln
  - Nutzergemeinde einbinden
  - Einheitliche und einfache Lizenzen



## Einschub: Position der Statistik im Open Data-Kontext

Seit dem Inkrafttreten der „Verordnung (EG) Nr. 223/2009 des Europäischen Parlamentes und des Rates vom 11.3.2009 über europäische Statistiken“ (EUStatVO) am 1.4.2009 besteht ein neuer Rechtsrahmen für die Entwicklung, Erstellung und Verbreitung amtlicher Statistiken. Nach Art. 2 EUStatVO sind die **Qualitätsnormen** des ESS für die Verbreitung amtlicher Statistik anzuwenden. Hiernach müssen die statistischen Daten **nutzerorientiert klar und verständlich präsentiert** und in **benutzerfreundlicher Weise verbreitet** werden und zusammen mit einschlägigen **Metadaten** und Erläuterungen entsprechend dem Grundsatz der **Unparteilichkeit** verfügbar und zugänglich sein. Hierzu sind moderne Informations- und Kommunikationstechnologie zu nutzen und die Qualitätsnormen des ESS anzuwenden.

## Einschub: Position der Statistik im Open Data-Kontext

Prinzipien	Regelungen in amtlicher Statistik
Vollständigkeit	Veröffentlichungsgebot: Gilt für Daten der amtlichen Statistik gemäß §1 BStatG und §2 LStatG Berlin /§3 BbgStatG Brandenburg
Primärquelle	Die Statistischen Ämter sind die Quelle der Daten. Einmal veröffentlichte Daten bleiben erhalten.
Zeitnah	Ergebnisse werden gemäß dem Veröffentlichungskalender so schnell wie möglich im Internet bereitgestellt.
Zugänglichkeit	Daten stehen allen grundsätzlich kostenlos im Internet zu Verfügung
maschinenlesbar / nicht proprietär	Formate: *.csv, *.xls etc.
nicht diskriminierend	Verfügbare Daten werden jedem zur gleichen Zeit zur Verfügung gestellt (geregelt durch Code of Practice - EU-Verordnung). D.h. auch keine Vorab-Bereitstellung für die Verwaltung oder die politische Ebene.
Lizenzfrei	Dieser Punkt ist klärungsbedürftig

## Phase 2: Regelbetrieb Datenportal und weiterer Ausbau

- Koalitionsvereinbarung nach Regierungswechsel 2011/12:  
„Die Koalition wird die Open-Data-Initiative des Landes fortsetzen und ausbauen. Dazu setzt sie sich für eine Prüfung der weitgehenden Offenlegung von öffentlichen Daten (z. B. Geoinformationsdaten) unter Wahrung des persönlichen Datenschutzes ein.“
- Einrichtung einer Arbeitsgruppe „Open Data Berlin“ mit den Themen:
  - Dateiformate und Metadaten
  - Lizenzen und Nutzungsbedingungen
  - Weiterbildung und Qualifizierung
  - Organisation und Grundsatz
- Überführung des Portals in den Regelbetrieb...

## Warum geht es nicht schneller voran?

### Politisch-institutionell

- Open Data ist auf die Arbeitsebene verlagert – ein CIO wurde nicht eingesetzt.
- Verwaltung handelt auf der Grundlage von Gesetzen, Verordnungen, Senatsbeschlüssen – eine entsprechende Grundlage wurde nicht geschaffen.

### Lizenzmodell: Rechtssicherheit für Datenbereitsteller und Nutzer

- **Forderung:** leicht verständlich, Rechtssicherheit gebend, netzweite Gültigkeit (beispielgebender Anspruch der CC-Lizenz)
- **CC-BY 3.0** nicht für Daten geeignet
  - Die Kurzfassung ist eingängig und plausibel (wirft aber einige Fragen auf), die Langfassung führt zu Unsicherheit und Ablehnung beim Datenbereitsteller.
- Vermesser **GeoNutzV** ohne Bezug auf CC und damit eine „lokale“ Regelung.
- **Datenlizenz Deutschland** ohne Bezug auf CC und damit eine „lokale“ Regelung.
- $\Sigma$ : Für die Daten eines Portals gelten verschiedene Regelungen?

## Warum geht es nicht schneller voran?

**Aus statistischer Sicht** (eine Meinungsäußerung) :

- Open Data darf **nicht** zu zusätzlichen Produkten und einem **zusätzlichen Aufwand führen**, wie: eigene Open Data-Läufe, Dokumente erstellen, in Katalog eintragen. Bei einer relevanten Quelle je Amt mag es noch angehen, nicht bei 300 Statistiken.
- Open Data muss in laufende **Veränderungsprozesse integriert** werden. Statistik bewegt sich von printförmigen Produkten zu Web-Datenbanken. AfS baut mit StatIS eine Datenbank auf, die auf **Mikrodaten** basiert. Näher an die „Rohdaten“ geht es aus Gründen der statistischen Geheimhaltung nicht. Es fehlt die Schnittstelle (API z.B.).
- Metadatenaustauschformat ist SDMX. Wie SDMX für Open Data nutzen?
- Wie verhält sich **Open Data** zu **INSPIRE**? INSPIRE/GDI bietet Dienste und Metadatenschemata für eine Programm-mit-Programm-Kommunikation.
- Open Data sollte keine Parallelstruktur zu INSPIRE/GDI werden. Ist ein anderes **kulturelles** und **Lizenzmodell**. csv-, xls-Downloads sind dagegen ein Rückschritt. Synergien nutzen!

## Warum geht es nicht schneller voran?

### Geschäftsprozesse

Es fehlen konzeptionelle Überlegungen zu den Geschäftsprozessen einer veränderten Datenkultur.

### Verwendung und Nutzen:

- Der Nutzen (gesellschaftlich, wirtschaftlich, verwaltungsinternes Moment) muss besser erkennbar werden.

### Nutzergemeinschaft:

- Bessere Vernetzung mit Nutzergemeinschaft. Der Umgang könnte fairer sein: nicht Datenbeamte etc. Nutzer sind auch wichtige Pressure Group (Ausfall der „Piraten“). Verwaltung müsste dazu in den sozialen Netzwerken präsent sein.

# Open Data-Angebot AfS

- Spezielle Seite für Open Data-Angebot.
- Wird aus einem bestehenden Datenpool versorgt, dennoch zusätzlicher Aufwand.

statistik Berlin Brandenburg

Statistiken Presse **Produkte** Datenbank Regionaldaten Datenerhebung Links

## Open Data

- Fläche und Gebiet
- Einwohner
- Mikrozensus
- Geometrien

<< zu Produkte

home → Produkte → Open Data

## Open Data

Hier finden Sie für **Open Data** aufbereitete Dateien in maschinenlesbarer Form für die weitere Nutzung durch die Öffentlichkeit, Wirtschaft, Schule und andere Institutionen. Unter „Statistiken“ sind xls-Dateien verfügbar.

Offene Daten werden ohne die Nutzung einschränkende Lizenzbestimmungen in maschinenlesbarer Form angeboten, um Innovationen, Transparenz und die Beteiligung von Bürgern zu unterstützen. Das Angebot befindet sich im Aufbau.

## Nutzungsbedingungen



Die Daten sind zur vereinfachten und transparenten Nutzung nach der Creative Commons Namensnennung 3.0 Deutschland lizenziert.  
(CC-BY 3.0 <http://creativecommons.org/licenses/by/3.0/de>)



→ Link zum Berliner Datenportal

→ "Berliner Open Data Day"

Über uns | Kontakt | Impressum |

© Amt für Statistik  
Berlin-Brandenburg 2013

# Open Data-Angebot AfS

- Download csv, kml, PDF (für Metadaten)

statistik Berlin Brandenburg

Statistiken

Presse

Produkte

Datenbank

Regionaldaten

Datenerhebung

Links

Open Data

- Fläche und Gebiet
- Einwohner
- Ortsteile
- Lebensweltlich orientierte Räume (LOR)
- Mikrozensus
- Geometrien

<< zu Produkte

home → Produkte → OpenData → Einwohner → LOR

## Lebensweltlich orientierte Räume

### Merkmale:

- LOR(447), Einwohner nach Geschlecht, Altersgruppen, Staatsangehörigkeit

### Quelle:

- Amt für Statistik Berlin-Brandenburg

### Beschreibung:

- Einwohnerregisterstatistik → PDF

### Download...

Einwohnerregisterstatistik -> csv-Datei -> Stand: 12/2011 alle Einwohner ▾

download

→ **BERLIN  
OPEN  
DATA**  
Link zum Berliner  
Datenportal

Über uns | Kontakt | Impressum

© Amt für Statistik  
Berlin-Brandenburg 2013

# Open Data-Angebot AfS

- StatS steht noch unter ©, könnte unter CC gestellt werden.

**statistik** Berlin Brandenburg

Datenbanken Tabellen Felder

TableVIEW ChartVIEW ColourVIEW

## Einwohnerregisterstatistik Berlin <sup>1</sup>

Wert: Summe  
 Felder: Region nach Stichtag und Migrationshintergrund

Stichtag	31.12.2011		
Migrationshintergrund	Deutsche ohne Migrationshintergrund	Deutsche mit Migrationshintergrund	Ausländer
Region			
080101 Schillerpromenade	48,4%	19,0%	32,6%
08010115 Hasenheide	63,0%	15,3%	21,7%
08010116 Wissmannstraße	47,8%	19,9%	32,2%
08010117 Schillerpromenade	52,6%	16,1%	31,3%
08010118 Silbersteinstraße	41,8%	23,0%	35,2%

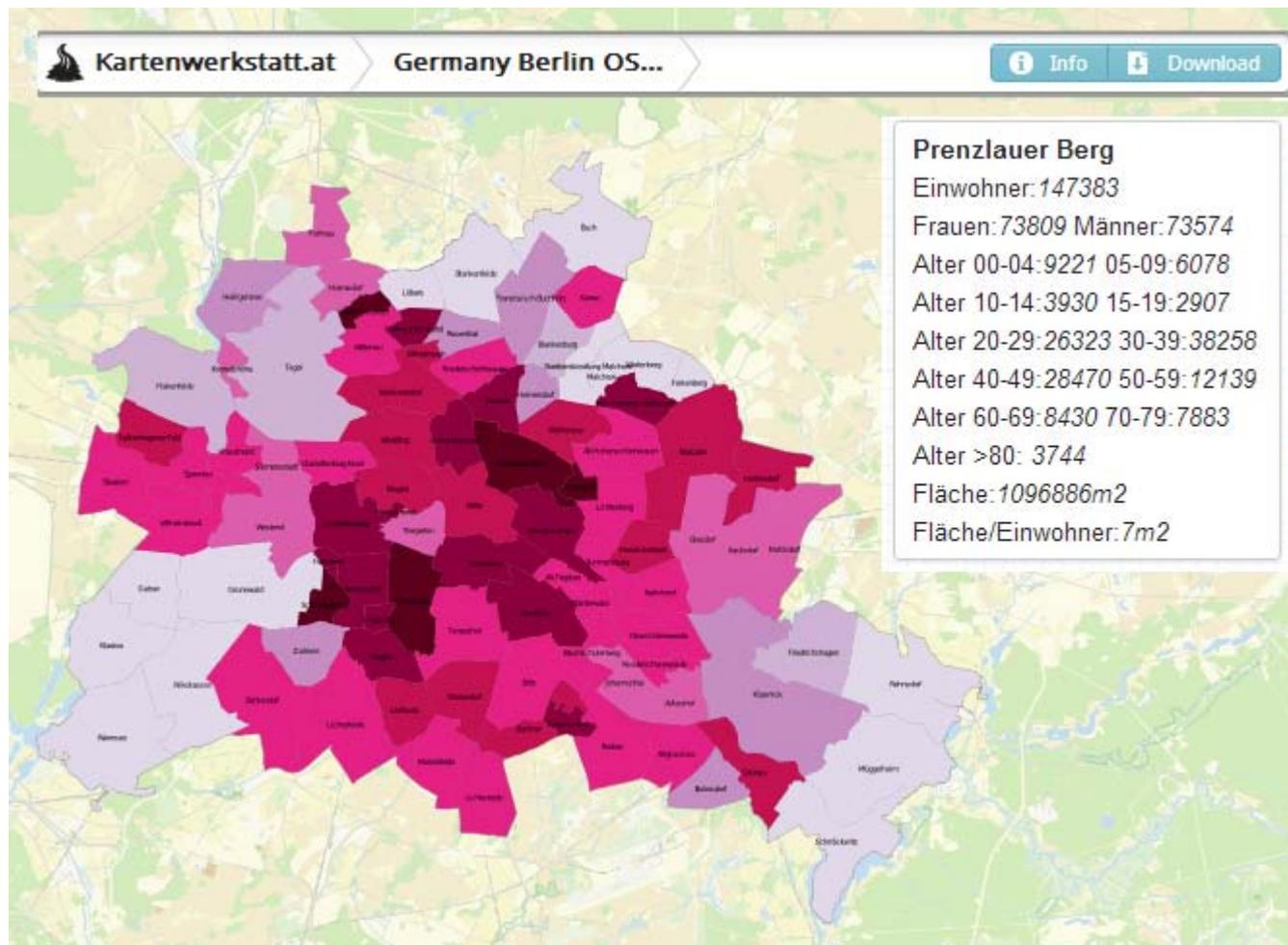
1 : Quelle: © Amt für Statistik Berlin-Brandenburg

Die Daten entstammen dem Einwohnerregister Berlins.  
 Kleinere Abweichung zu den Angaben der Bevölkerungsstatistik sind methodisch bedingt.

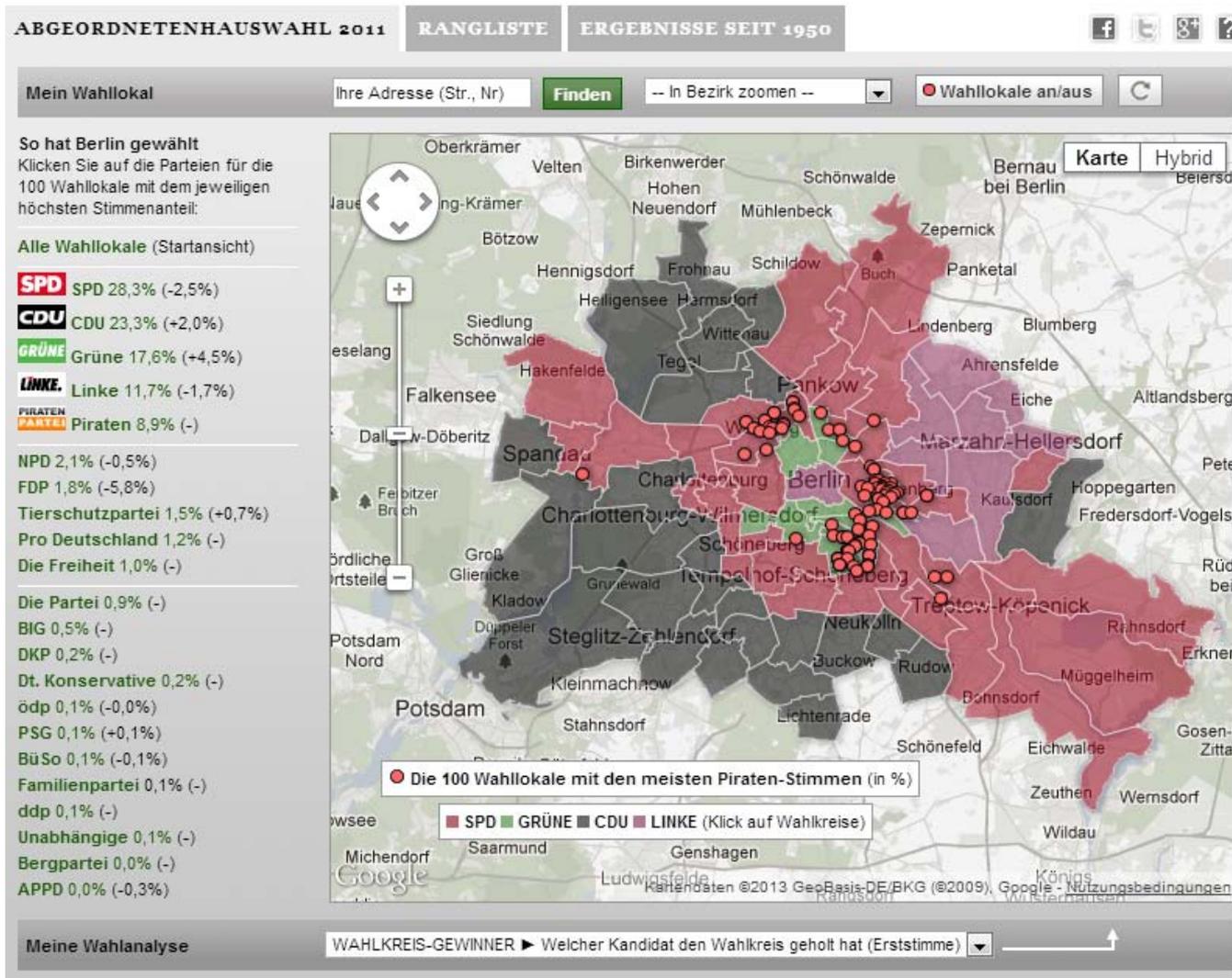
Für nähere Informationen zur Gliederung der Bezirke nach Lebensweltlich orientierten Räumen klicken

Tabelle:

## Beispiele für Verwendung offener Statistikdaten



# Beispiele für Verwendung offener Statistikdaten



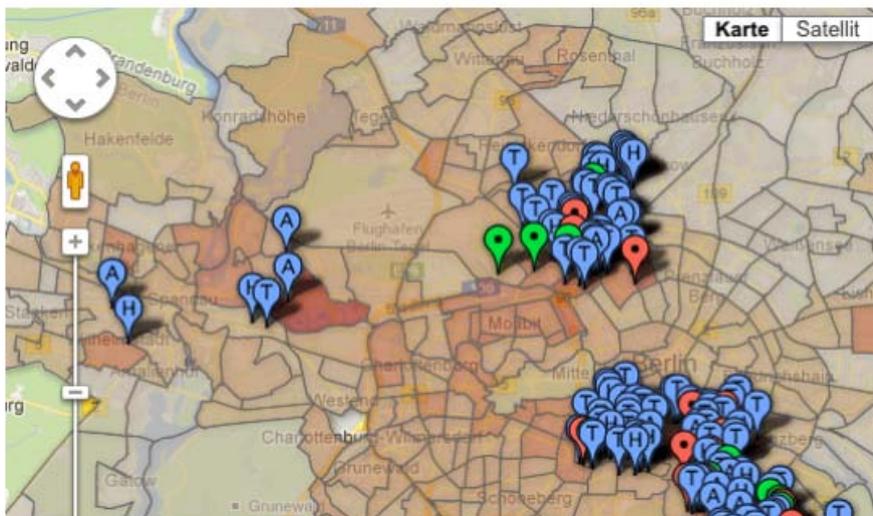
Quelle: Endgültiges amtliches Ergebnis, Stand: 06. Oktober 2011, Amt für Statistik Berlin-Brandenburg  
Julius Tröger, André Pätzold, Berliner Morgenpost, 2011

Berliner Morgenpost

# Beispiele für Verwendung offener Statistikdaten: Datenjournalismus

## “Vermisst”-Aktion beschränkt sich auf Klischee-Bezirke

VON Michael Hörz 30. NOVEMBER 2012 UM 14:32 UHR



Berliner Stadtteile mit Bewohnern mit Migrationshintergrund und Standorte der BMI-Aktion "Vermisst"  
Quelle: Frag den Staat

Bei der “Vermisst”-Aktion, die auf die Gefahr hinweisen will, dass junge Menschen in den islamischen Extremismus abdriften könnten, haben sich das Bundesinnenministerium und/oder die beteiligten Werbeagenturen offensichtlich von Klischees leiten lassen. Diese Vermutung stützen

...

**Nachtrag:** Die Datenquellen zur Berliner Einwohnerregisterstatistik hatte ich in der Ur-Version des Textes verlinkt, in der redigierten Fassung nicht mehr. Hier die [Einwohnerregisterstatistik](#), Stichtag 31.12.2011 und die [Shapefiles der Planungsräume](#).

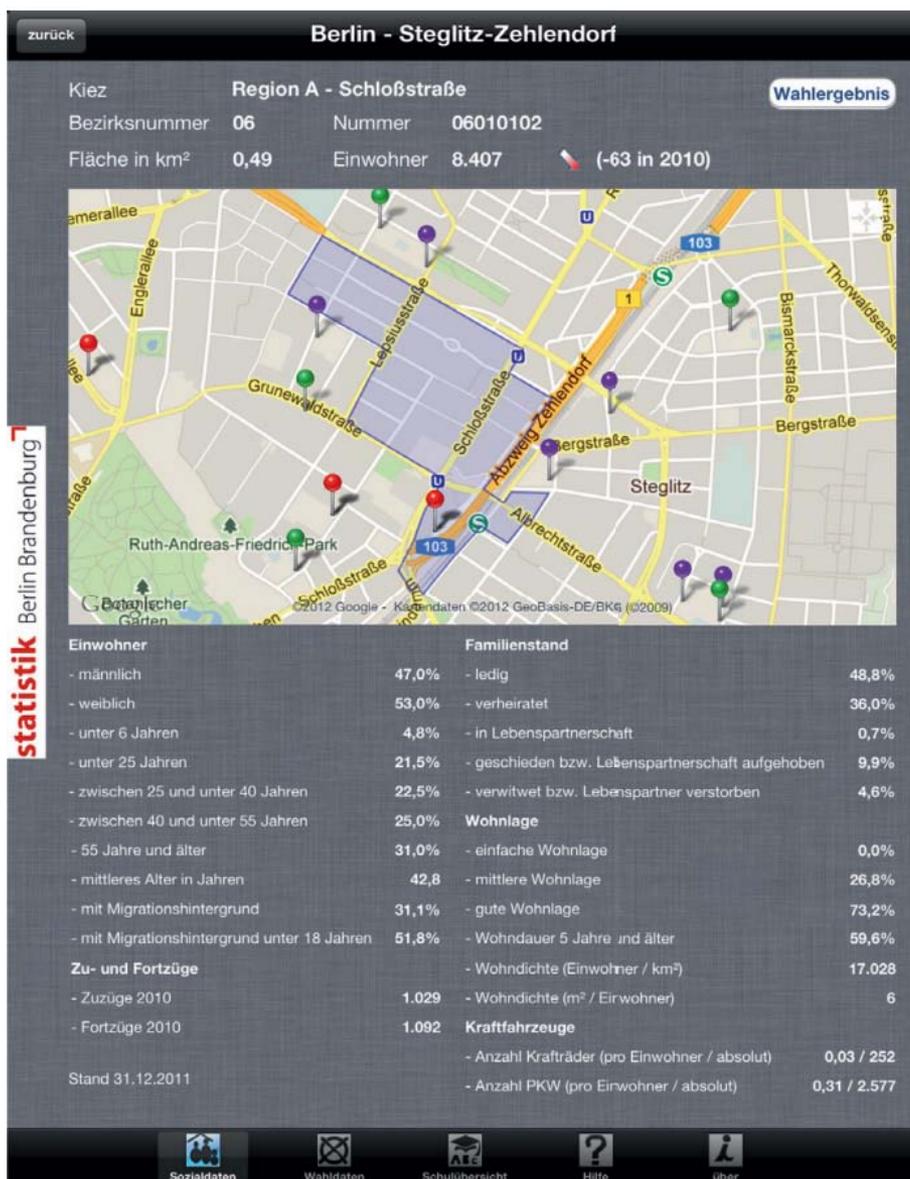
Quelle:



Korrektur einer vergessenen CC Namensnennung

# Beispiele für Verwendung Statistikdaten

KiezDaten-App (iOS) des AfS



## Nutzen und Risiken

- Zeitgemäße Bereitstellung weiterverarbeitbarer Daten: **Daten für Anwendungen**, nicht nur für Leser.
- Änderung der “**Datenkultur**” innerhalb der Verwaltung (“digitale Verwaltung”).
- Bessere Datenverfügbarkeit von Verwaltungsdaten für Aufgaben der Statistik in der Gemeinde.
- Einbringen unserer Kenntnisse und Erfahrungen bei der **Q-Sicherung**.
- Verlust der Sichtbarkeit und der Zuordnung von erbrachten Leistungen.
- Eigene “Datenveredelung” wird anspruchsvoller sein müssen.
- Bestreiten des Rechts auf eigene Datenveredelung durch externe Verwender dieser Daten.

# Open Data Census

 Open Data Census Dashboard
 [Home](#)
[Census](#)
[Catalogs](#)
 Open Knowledge Foundation

## Data Catalogs

Total datastores in catalog **283**

