

Öffentlich verfügbare Datenquellen

<https://www.quora.com/Where-can-I-find-large-datasets-open-to-the-public>

Name	Beschreibung	URL	Protokoll + Format	Beispieldatenelement
Meetup.com RSVP Stream	Zusagen zu einem öffentlichem Event	Live-View des Datenstreams: http://meetup.github.io/stream/rsvpTicker/ Doku: http://www.meetup.com/meetup-api/docs/stream/2/rsvps/	JSON über Websocket oder Chunked HTTP	<pre>{"venue":{"venue_name":"The Salon Lounge, Blythswood Square Hotel.,"lon":-4.262657,"lat":55.864048,"venue_id":23858117},"visibility":"public","response":"yes","guests":0,"member":{"member_id":192306204,"member_name":"Victoria "},"rsvp_id":1577610380,"mtime":1445960413000,"event":{"event_name":"It Girl Glasgow ~ Glamorous Networking For Positive Women","event_id":"224631531","time":1445972400000,"event_url":"http://www.meetup.com/It-Girl-Glasgow/events/224631531/"},"group":{"group_topics":[{"urlkey":"women","topic_name":"Women's Social"}, {"urlkey":"professional-networking","topic_name":"Professional Networking"}, {"urlkey":"self-improvement","topic_name":"Self-Improvement"}, {"urlkey":"lawofattraction","topic_name":"Law of Attraction"}, {"urlkey":"lifetransform","topic_name":"Life Transformation"}, {"urlkey":"women-entrepreneurs","topic_name":"Women Entrepreneurs"}, {"urlkey":"personal-development","topic_name":"Personal Development"}, {"urlkey":"womens-empowerment","topic_name":"Women's Empowerment"}, {"urlkey":"life-coaching","topic_name":"Life Coaching"}, {"urlkey":"entrepreneurship","topic_name":"Entrepreneurship"}, {"urlkey":"nlp-coaching","topic_name":"NLP Coaching"}, {"urlkey":"self-empowerment","topic_name":"Self-Empowerment"}, {"urlkey":"nlp-neurolinguistic-programming","topic_name":"NLP (Neuro-Linguistic Programming)"}, {"urlkey":"hypnosis-and-hypnotherapy","topic_name":"Hypnotherapy"}, {"urlkey":"boss-ladies","topic_name":"Boss Ladies"}]}, "group_city":"Renfrew","group_country":"gb","group_id":17310722,"group_name":"It Girl Glasgow - Glamorous Events For Positive Women","group_lon":-4.41,"group_urlname":"It-Girl-Glasgow","group_lat":55.88}}</pre>

Transport for London Data Feed	ÖPNV-Daten • Journey Planning (current and future) • Status (current and future) • Disruptions (current) and Planned works (future) • Arrival/departure predictions (instant and web sockets) • Timetables • Embarkation points and facilities • Routes and lines (topology and geographical) • Fares	https://tfl.gov.uk/info-for/open-data-users/data-feeds?intcmp=29422 https://api.tfl.gov.uk/		
CLEF-Newsreel	Aufrufe von Nachrichtenartikeln und Aufforderungen, Empfehlungen zu geben	http://www.clef-newsreel.org/	JSON über HTTP-Post	
NWS Public Alerts	Wetterbenachrichtigungen	http://alerts.weather.gov/	XML/CAP and ATOM Format	
Yahoo Query Language	YQL Web Service	https://developer.yahoo.com/yql/guide/	XML/JSON	https://developer.yahoo.com/yql/console/
Kaggle	Public datasets hosted by Kaggle	https://www.kaggle.com/datasets		

Wikidata	Free linked database of wikipedia data	https://www.wikidata.org/wiki/Wikidata:Main_Page https://www.mediawiki.org/wiki/Wikibase/API	JSON/XML	<pre> { "batchcomplete": "", "query": { "pages": { "214": { "pageid": 214, "ns": 0, "title": "Q84", "terms": { "description": ["capital city of England and the United Kingdom", "capital city of England and the United Kingdom"], "alias": ["London, England", "London, UK", "London, United Kingdom", "London, England", "London, UK", "London, United Kingdom"], "label": ["London"] } } } } } </pre>
Intel Lab Data	Sensor measurements from 54 sensors deployed in the Intel Berkeley Research lab	http://db.csail.mit.edu/labdata/labdata.html	CSV	
UC Irvine Machine Learning Repository	Repository with 335 datasets for the machine learning community	https://archive.ics.uci.edu/ml/	CSV	
Amazon	Different datasets hosted by Amazon AWS	https://aws.amazon.com/public-data-sets/		
Google	Different datasets hosted by Google	http://www.google.com/publicdata/directory#!		
KDD Cup	Datasets used for the annual Data Mining and Knowledge Discovery competition organized by ACM Special Interest Group on Knowledge Discovery and Data Mining	http://www.kdd.org/kdd-cup	CSV	
MarineCadastre	AIS data from the US coast as GPS trajectories	http://marinecadastre.gov/ais/	GDB, can be exported to CSV via QGIS (http://www.qgis.org/de/site/)	X, Y, SOG, COG, Heading, ROT, BaseDateTime, Status, VoyageID, MMSI, ReceiverType, ReceiverID -177.234823, 60.6791, 0.1000000001490116, 328.299987792969, 511, 0, 2014/01/01 01:44:46, 9, 1, 367897740, D, 08MN

GeoLife	Peoples trajectory data from social networks (GPS measurements of their movement)	https://www.microsoft.com/en-us/research/project/geoife-building-social-networks-using-human-location-history/ https://www.microsoft.com/en-us/download/details.aspx?id=52367	CSV	39.984094, 116.319236, 0, 492, 39744.2451967593, 2008-10-23, 05:53:05 39.984198, 116.319322, 0, 492, 39744.2452083333, 2008-10-23, 05:53:06
Udacity Self Driving Car	223GB of driving data with location (lat/long), gear, break, throttle, steering angle, speed and image	https://github.com/udacity/self-driving-car https://medium.com/udacity/open-sourcing-223gb-of-mountain-view-driving-data-f6b5593fbfa5#pe7j0hi8f https://www.udacity.com/self-driving-car	CSV and images	Latitude, Longitude, Gear, Brake, Throttle, Steering Angle, Speed, FileName 37.399960, -122.131840, 4, 0.147433, 0.307836, 0.005236, 10.150000, images/1475187707065512506.png 37.399813, -122.132192, 4, 0.213535, 0.149950, 0.024435, 0.000000, images/1475187679161015902.png 37.398688, -122.134251, 4, 0.147890, 0.285496, 0.144862, 6.222222, images/1475187468081761839.png
GDELT	Geo-Referenced data that includes social happenings such as protests, violence reports, etc. Newest version updates every 15 minutes.	http://gdeproject.org/	CSV	
Sloan Digital Sky Survey	Data from Apache Point Observatory, New Mexico	http://www.sdss.org		
National Renewable Energy Laboratory	Datasets from wind power plants in North America	http://www.nrel.gov/electricity/transmission/western_wind_disclaimer.html		
OpenEI	Open Energy Information: buildings, geothermal, hydrogen, smart grid, solar, utilities, water, wind	http://openei.org		
T-Drive trajectory data	Data from taxis trajectories	https://www.microsoft.com/en-us/research/publication/t-drive-trajectory-data-sample/	CSV	39, 2008-02-02 13:37:30, 116.29369, 39.92272 39, 2008-02-02 13:40:17, 116.28015, 39.92321 39, 2008-02-02 13:45:17, 116.28065, 39.9233 39, 2008-02-02 13:45:17, 116.28065, 39.9233 39, 2008-02-02 13:49:15, 116.28012, 39.92327