*Merseburg University of Applied Sciences; Institute for Applied Informatics e. V., Leipzig; CISS TDI GmbH, Sinzig; Chemnitz University of Technology*

# MCLIENT – A WEB TOOLKIT FOR OPEN DATA PUBLISHING

Lisa Wenige, Claus Stadler, Christopher W. Frank, Michael Martin, Richard Figura

**Abstract:** The mobility data client (mCLIENT) is a web application designed to simplify the publication of datasets to Open Data portals by providing automated quality control, metadata enhancement, and seamless publication features. Initially implemented for General Transit Feed Specification (GTFS) files, the tool's methodology is broadly applicable to various data types and formats, aiding sustainable urban planning and traffic management. Despite the rise of Open Data portals, many regions lack the technical tools needed for effective data publishing. Existing platforms like CKAN and DKAN offer partial solutions but miss crucial automation features. mCLIENT addresses these gaps by automating workflows to improve data quality and reduce manual effort. Expert surveys highlight the urgent need for such tools, and mCLIENT's adaptable approach promises advancements within the mobility sector as well as across various data domains.

**Keywords:** Open Data, mobility data, data quality, metadata management, dataset publication, data deployment

## MCLIENT – EIN WEB-TOOLKIT ZUR VERÖFFENTLICHUNG OFFENER DATEN

**Zusammenfassung:** Der Mobility Data Client (mCLIENT) ist eine Webanwendung, die entwickelt wurde, um die Veröffentlichung von Datensätzen auf Open-Data-Portalen zu vereinfachen, indem sie automatisierte Qualitätskontrollen sowie Funktionen zur Metadatenanreicherung und Publikation von Datensammlungen bereitstellt. Die Anwendung verarbeitet Dateien, die der General Transit Feed Specification (GTFS) entsprechen, demonstriert aber eine Methodik, die auf verschiedene Datentypen und -formate anwendbar ist und damit ein nachhaltiges Verkehrsmanagement unterstützen kann. Trotz der hohen Verfügbarkeit von Open-Data-Portalen fehlt es vielerorts an den technischen Werkzeugen, die für eine effektive Datenveröffentlichung erforderlich sind. Bestehende Plattformen, wie CKAN und DKAN, bieten teilweise Lösungen, es fehlen jedoch wichtige Automatisierungsfunktionen. mCLIENT schließt diese Lücken, indem es Arbeitsabläufe automatisiert, um die Datenqualität zu verbessern und den manuellen Aufwand zu reduzieren. Literaturanalysen und Expertenumfragen betonen den dringenden Bedarf an solchen Werkzeugen. Der anpassungsfähige Ansatz von mCLIENT verspricht verbesserte Publikationsprozesse für Mobilitätsdaten.

**Schlüsselwörter:** Open Data, Mobilitätsdaten, Datenqualität, Metadatenmanagement, Datenpublikation, Datendeployment

## 1 INTRODUCTION

We present the mobility data client (mCLIENT) – a web application that supports publishing datasets to Open Data portals. The tool handles common problems of dataset deployment by providing features such as automated quality control and metadata enhancement as well as seamless publication of datasets to remote platforms. In its current state, the tool supports the deployment of files formatted according to the General Transit Feed Specification (GTFS), a standardized format for public transportation schedules and related geographic data. The tool's methodology is versatile and applicable to diverse data types and formats, potentially supporting a wide range of mobility data.

Enhanced availability of mobility data can facilitate better linking and integration of public transportation routes, thereby supporting more environmentally friendly and sustainable urban and traffic planning efforts. Mobility data refers to information that describes people's movements between locations as well as their selected mode of transportation. Access to mobility data is therefore essential for epidemic control, traffic management, urban planning and the development of effective mobility services (Zhao et al. 2016).

Data platforms play a crucial role by providing relevant datasets for monitoring traffic services. Recent years have seen an increase in the existence of Open Data portals. Notably, dataportals.org curates a comprehensive list of global Open Data platforms. Since its inception in 2011, the number of publicly accessible data portals has doubled, totaling nearly 600 portals worldwide as of 2024[1].

This progression is intricately tied to the recognition among global political decision-makers of the pivotal role played by

[1] https://github.com/okfn/dataportals.org


DOI: 10.14627/gis.Science.2024.3.3

Authors

Prof. Dr. Lisa Wenige
Merseburg University of Applied Sciences
Department of Engineering and Natural Sciences
Eberhard-Leibnitz-Str. 2
D-06217 Merseburg
E: lisa.wenige@hs-merseburg.de

Claus Stadler
Institute for Applied Informatics e. V.
Goerdelerring 9
D-04109 Leipzig
E: stadler@infai.org

Dr. Christopher W. Frank
Dr. Richard Figura
CISS TDI GmbH
Barbarossastraße 36a
D-53489 Sinzig
E: c.frank@ciss.de
r.figura@ciss.de

Prof. Dr. Michael Martin
Chemnitz University of Technology
Faculty of Computer Science
Professorship of Data Management
Straße der Nationen 62
D-09111 Chemnitz
E: dm@informatik.tu-chemnitz.de

Open Data in enhancing transparency and optimizing public services. Further reinforcement for such advancements comes from corresponding legislation, manifesting both internationally and nationally. As a case in point, mobility data stands out as one of the six categories of high-value datasets (HVD) that public institutions in Europe are mandated to publish (European Commission 2023). While this dedicated legal framework facilitates increased data availability, in countries such as Germany still only a fraction of administrative units at NUTS level 2 or lower maintain their own Open Data platforms (Wenige et al. 2021).

One of the reasons for the limited coverage is the lack of technical tools assisting data providers in preparing datasets for publishing. While there already exist powerful open source software for data platforms, such as CKAN or DKAN, their support for automatic quality control and metadata enhancement is still rather limited. Moreover, only a fraction of globally prevalent portals utilizes these software platforms. Some instead opt for a custom solution that, in turn, does not provide a public API for harvesting (pull principle) or publishing data (push principle) (Hinz & Bill 2021, Neumaier et al. 2016, Wenige et al. 2021). Data providers not connected to a platform face significant challenges in making their data publicly available. Thus, there is a need to develop efficient software tools that support the publication process. Unfortunately, such tools are not universally accessible, posing challenges to implementing legal standards and serving as a hurdle to widespread Open Data publishing.

This article demonstrates how existing web services can be merged into a unified user interface thereby mitigating typical problems of Open Data publishing. The mCLIENT application can help data-providing institutions by overcoming technical barriers and automating workflows commonly occurring in data publishing.

## 2 LITERATURE REVIEW: TECHNICAL BARRIERS TO OPEN DATA PUBLISHING

The literature mentions multiple barriers with regard to Open Data publishing from the perspective of data publishers. Frequently stated challenges include a lack of political support, constraints in resources such as technical infrastructure and expertise, along with concerns related to data quality and privacy issues (McBride et al. 2018). With regard to effective IT design, the aspects of data quality and missing IT support have been often mentioned in empirical studies conducted among data suppliers. These aspects will be explained in more detail in the following sections.

### 2.1 DATA HANDLING

Researchers have studied problems of Open Data publishing both with qualitative as well as quantitative methods identifying major barriers in the areas of dataset quality, metadata quality as well as metadata integration.

**Dataset Quality:** Moyano et al. as well as Nikiforova demonstrated that the quality of many published datasets is significantly compromised (Moyano et al. 2017, Nikiforova 2020). Expert interviews conducted with senior government managers in Ireland and the Netherlands (Barry & Bannister 2014, Conradie & Choenni 2014) further unveiled substantial concerns from policy makers regarding data publishing. Decision makers specifically emphasized the potential for misinterpretation of data and the existence of errors in datasets. Janssen et al. reported similar findings, with interviewees underscoring data quality issues related to the accuracy and completeness of information, as well as the presence of obsolete and non-valid data, which in turn prevents a broader use of the data (Janssen et al. 2012). This finding was also confirmed in our expert interviews (see section 3): Data providers are often reluctant to publish datasets because the existing data quality does not meet minimum standards.

**Metadata Quality:** Additionally, experts mentioned that users might struggle to identify relevant collections in the sea of existing datasets (Janssen et al. 2012). The latter claim is also supported by findings from quantitative analyses conducted with regard to the metadata quality in Open Data portals. For instance, in our analysis of the German Open Data landscape we found out that the informativeness of used keywords, measured using Shannon Information Content, is relatively low. Many datasets in data catalogs use non-meaningful keywords, making it challenging for users to locate specific collections. Similarly, other dataset fields such as title or description are often so generic that identically named fields occur frequently, even though the compared datasets are not genuine duplicates (Wenige et al. 2021). In addition, the metadata-based worldwide Open Data Benchmark by Neumaier et al. confirms that many data fields, such as license or contact information, are either not filled or contain incomplete information (Neumaier et al. 2016). Hinz & Bill state a high degree of technical and syntactical heterogeneity in data portals, affecting interoperability. Non-standardized values in metadata fields, such as licenses, create complexities, as different identifiers are used across portals. Thus, data discovery challenges arise which further build barriers for

datasets to be found in upstream national and international portals (Hinz & Bill 2021).

Metadata Silos: Another area of untapped potential is the integration with the web of data. Although the DCAT metadata standard and its European application profile, DCAT-AP, are designed as machine-readable semantic formats for seamless integration into existing knowledge graphs, this potential remains largely underutilized. In our analysis of the German Open Data landscape, examining crawled catalogs from over 60 portals, we primarily identified the standard namespaces for data catalogs (DCAT, spdx, adms) and the semantic web standards (RDF, OWL). Additionally, we observed vocabularies aligned with the DCAT standard for storing contact information (foaf) and keywords (dct). Figure 1 shows the distribution of namespaces in the German Open Data landscape listing all namespaces that occurred more than 100 times in the catalog. No other namespaces were identified (Wenige et al. 2021). This is surprising, given that the DCAT standard explicitly supports the integration of additional vocabularies (e. g. SKOS) and presents a prime opportunity to link keywords, individuals, or organizations with other knowledge graph repositories like Wikidata or DBpedia[2].

## 2.2 SERVICE INTEGRATION AND AUTOMATION

Among the various challenges for Open Data is the lack of sufficient tool support during the publishing process. This results in many stakeholders, who manage relevant data, choosing not to publish it due to the perceived high associated effort (Janssen et al. 2012). The publishing process usually entails a set of subtasks which are currently only partially automated and can often not yet be processed within an integrated workflow.

The collection of subtasks involves checking dataset quality, creating metadata, verifying metadata quality, and deploying or publishing the data on an Open Data portal (see Figure 2). Table 1 lists the state of the art in terms of tool support for dataset publications and outlines current implications for the publishing process that result from a lack of service integration in this area.

Related Datasets Crawling: Employees tasked with managing pertinent data in public institutions and publishing it for the first time also face the challenge of initially reviewing comparable data on portals of similar platforms. This task is often carried out manually by searching on other portals. Additionally, there is the option to utilize services such as the "sample data catalog" ("Musterdatenkatalog") which gives an overview of available datasets in German municipalities (Wiedemann & Bürger 2021). However, in its current state this catalog is not provided in the DCAT format. Automated crawling of remote data portals can be facilitated through harvesting via a CKAN or DKAN instance. For this purpose, either local CKAN or DKAN instances must be available, which would allow initiating such crawling via a user or a command-line interface. However, the harvested metadata is imported into the portal which may not be desired for every use case.

Data Handling: Some tools already provide options for automated quality checking, such as CSV Validator, Frictionless Tables or the Canonical GTFS Schedule Validator. While the two former tools operate on CSV files of any kind, the latter one explicitly addresses quality control for mobility data. The GTFS Validator offers machine-readable JSON reports. However, error messaging in RDF formats would even be more beneficial to allow for seamless metadata shipping alongside the DCAT file. Additionally, it would be desirable to integrate the quality tool into the general publication process. In the field of metadata generation, there are similar gaps in the state of the art. In current data portals, data editors rely on manual user inputs of metadata. However, it would be helpful if annotation tools, such as DBpedia Spotlight (Mendes et al. 2011) could be integrated into the publication process in such a way that metadata descriptions are automatically enriched with appropriate keywords and thus link the dataset to the web of data. Previous research on the effectiveness of DBpedia Spotlight, which evaluated its performance using a manually

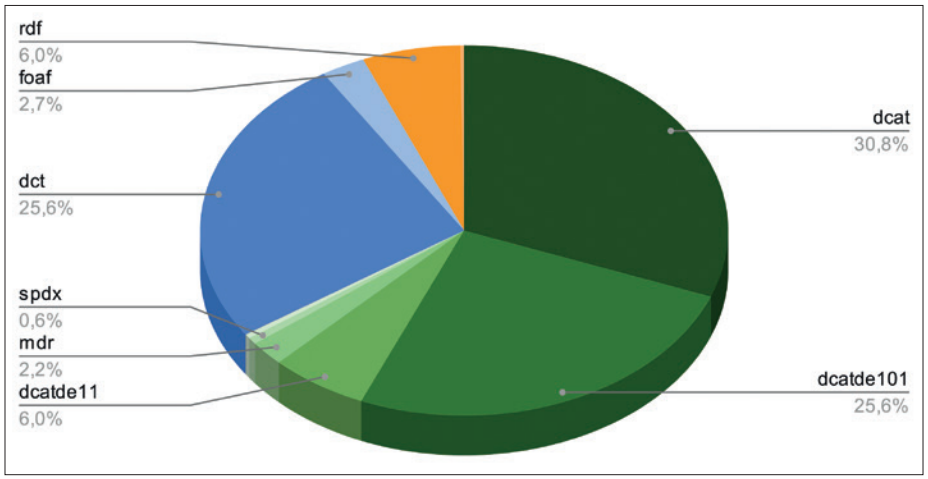


*Figure 1: Namespaces in the German Open Data landscape (The namespace distribution was obtained from querying the crawled data catalog: https://github.com/mclient-project/crawling)*

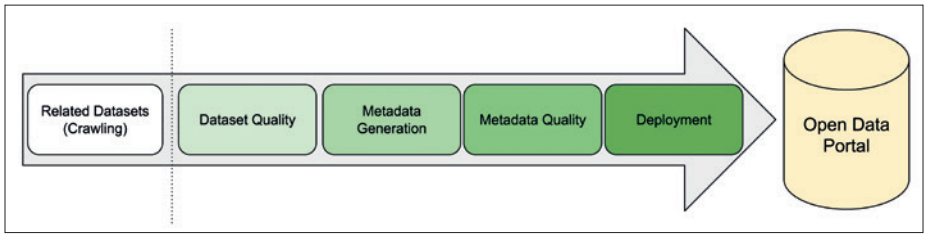


*Figure 2: Typical data publication process*

[2] https://lov.linkeddata.es/dataset/lov/vocabs/dcat

| Subtask | | Tool Support | Implications for Data Publishing |
|---|---|---|---|
| Related Datasets Crawling | | Plugins CKAN-harvest[1] & DKAN harvest[2] | • A data portal must be available<br>• Corresponding metadata is automatically imported into the portal which may not be desired |
| | | Sample data catalog[3] | • No DCAT metadata (CSV only) |
| Data Handling | Dataset Quality | CSV Validator[4] Frictionless Tables[5] | • Limited machine-readability of error reports |
| | | Canonical GTFS Schedule Validator[6] | • Non-RDF error reports<br>• Needs integration in publishing toolchains |
| | Metadata Generation | Metadata editors (e.g., in admin sections of CKAN or DKAN-based data portals) | • Manual metadata generation required<br>• Limited prevalence of DCAT support |
| | | DBPedia Spotlight | • Needs integration in publishing toolchains |
| | Metadata Quality | ITB DCAT-AP Validator[7], ITB DCAT-AP.de Validator[8] | Needs integration in publishing toolchains |
| Deployment | | CKAN API[9] / DKAN API[10], ckanapi CLI[11] | Needs integration in publishing toolchains |

*Table 1: Current tool support for data publication*

[1] https://github.com/ckan/ckanext-harvest
[2] https://dkan.readthedocs.io/en/latest/components/dkan_harvest.html
[3] https://www.bertelsmann-stiftung.de/de/unsere-projekte/smart-country/musterdatenkatalog
[4] https://github.com/digital-preservation/csv-validator
[5] https://repository.frictionlessdata.io/docs/getting-started.html#usage
[6] https://github.com/MobilityData/gtfs-validator
[7] https://www.itb.ec.europa.eu/shacl/dcat-ap/upload
[8] https://www.itb.ec.europa.eu/shacl/dcat-ap.de/upload
[9] https://docs.ckan.org/en/2.10/api/
[10] https://dkan.readthedocs.io/en/latest/user-guide/guide_dataset.html
[11] https://github.com/ckan/ckanapi

annotated corpus of mobility dataset descriptions, demonstrated a fairly good performance in terms of annotation accuracy (Wenige et al. 2020).

Additionally, not all portals provide API access to their metadata in DCAT format, once dataset descriptions have been entered (Wenige et al. 2021). This issue complicates the interoperability and machine-readability of the data, which may be necessary for subsequent workflow steps, such as checking for DCAT conformity. The Interoperability Test Bed (ITB), a service offered by the European Commission, is designed to streamline the conformance testing of IT systems and has been configured to automatically validate DCAT-AP compliance of metadata descriptions based on SHACL shapes. It is accessible through a web interface and ideally should be integrated into an automated Open Data quality workflow.

**Deployment:** At the end of a processing chain, the data collection is uploaded to the data portal, including the associated metadata. To perform this automatically, an API connection is required. Both CKAN and DKAN have such interfaces. Additionally, for the CKAN portal software, a command-line interface is also provided to support the deployment process. This existing tool support should be utilized in an automated pipeline for publishing data collections.

### 2.3 USER INTERFACES

The lack of appealing portal interfaces for searching geospatial and mobility-related data has already been noted in the literature (Bill et al. 2018, Hinz & Bill 2018). This gap also affects the data provision process in the backend of data portals. Empirical analyses confirm that missing technical resources or expertise within data-providing institutions are one of the most common obstacles to the publication of Open Data collections (McBride et al. 2018). While the general publication process of data collections in CKAN and DKAN instances is already supported by corresponding web interfaces and can be carried out without major usability issues (Akyürek et al. 2018), there is a lack of integrated user interfaces that adequately support automated quality checks and deployment to downstream data portals (see Table 1).

These steps can be conducted from the CKAN or DKAN interface by providing a suitable plugin for these systems. However, it must be considered that only about half to two-thirds of available data portals are operated with this software (Wenige et al. 2021). Therefore, it is advisable to aim for standalone interfaces that graphically support the publication process and can be seamlessly connected to CKAN or DKAN but also operate independently of these portal systems.

## 3 EXPERT SURVEY AND REQUIREMENTS ANALYSIS

To refine the insights derived from the literature on mobility data, we conducted a qualitative survey with six experts from Germany. Two of these experts represented transport associations, while the remaining four were operators of Open Data portals at both the municipal and federal levels. Some of these portals also serve as data providers for downstream aggregator portals. The data pro-

viders mostly focus on the publication of time-table and traffic related datasets such as GTFS files. The experts were asked about data provision and desirable improvements. The survey revealed the following steps of a typical publishing process:

1. If the data already exists in an established format, the raw data is first converted into the respective standard format.
2. In a few cases, particularly when a standard format is established, a quality check is subsequently conducted by the provider.
3. Metadata for the relevant dataset is collected and provided in a predominantly manual fashion.

The experts interviewed expressed a desire for a more automated workflow for this process. The specific filling of metadata fields is subject to individual discretion, resulting in a high manual effort for each new dataset. Additionally, there is a lack of automated completeness and plausibility checks. From the perspective of the portal operators, there is also a strong desire for improvements in the content quality of the provided datasets, as a significant portion of the data currently provided is of such low quality that it cannot be published. In summary, it can be concluded that the problems identified in the literature review regarding the data publication process have been confirmed by the expert survey in the field of mobility data.

## 4 MCLIENT

Based on the requirements as identified by the literature review and the mobility expert interviews, a reference architecture was created. It includes the essential components of the prototype as well as the intended interfaces to external web services for (meta-)data quality assurance (CISS Quality Assurance Center – CISS QuAC; ITB DCAT-AP Validator) and metadata enrichment (DBpedia Spotlight). Figure 3 illustrates the schematic layout of the mCLIENT demonstrator's reference architecture. The mCLIENT demonstrator is part of the dcat-suite[3] software library which handles data management for DCAT datasets. It also uses the Apache Vaadin framework for the web interface implementation[4].

To assess the system's operational effectiveness, we performed functional tests using the publication of GTFS data as a representative use case. The following sections demonstrate the functionality of our system's components using sample datasets.

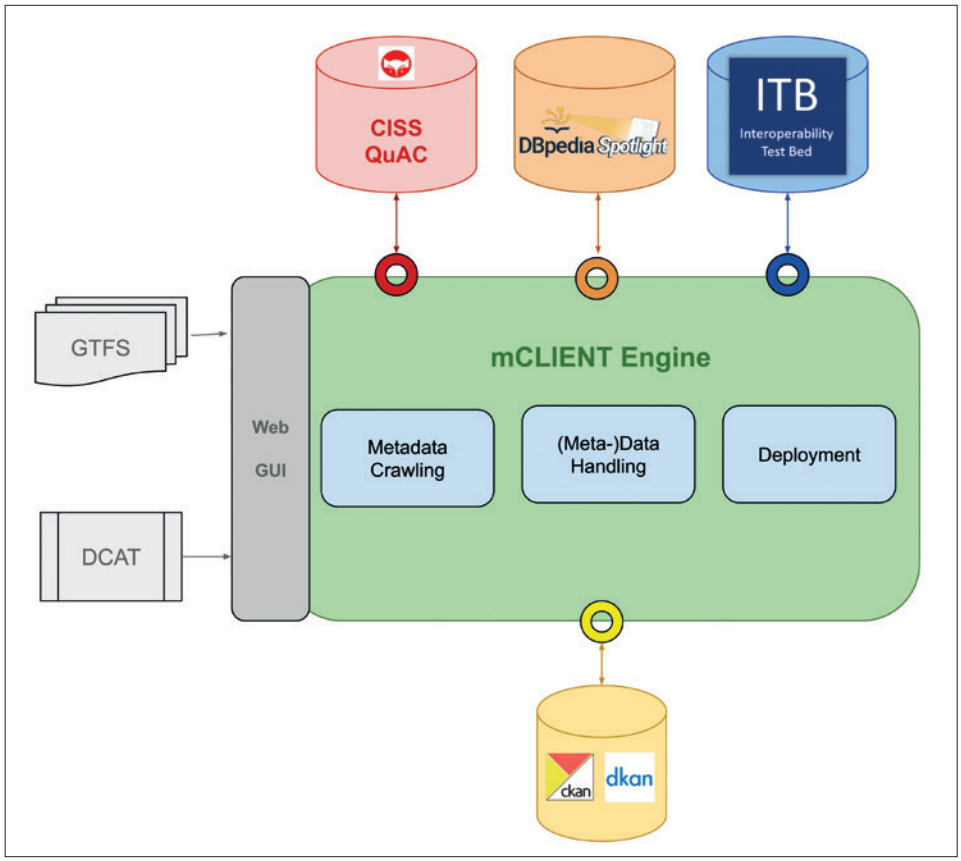


*Figure 3: mCLIENT architecture*

### 4.1 METADATA CRAWLING

As one of the key components of the mCLIENT demonstrator, the crawling component takes care of metadata extraction of remote Open Data portals. Thus, descriptions of datasets from portals that use a CKAN or DKAN interface can be obtained. The process involves systematically querying JSON metadata provided by these portals and converting it into DCAT-AP compliant metadata. To use this component, the URL of the relevant Open Data portal must be specified, along with the type of interface referring to CKAN or DKAN APIs, respectively. This distinction is important because data querying methods differ between these APIs, and the differences are accounted for in the software. With the crawling component DCAT-AP descriptions can be obtained even if the portal only provides metadata descriptions in JSON format. For our use case, we tested crawling on the DKAN-based Open Data portal of the city of Mühlheim (Ruhr), which currently does not provide a DCAT-AP interface[5]. Figure 4 illustrates the web interface of the component and shows a successfully completed crawling process.

### 4.2 DATA HANDLING

Quality assurance of data collections is a central part of the publication process. In response to transit agency representatives

[3] https://github.com/SmartDataAnalytics/dcat-suite

[4] https://vaadin.com

[5] https://geo.muelheim-ruhr.de/open-data/13819

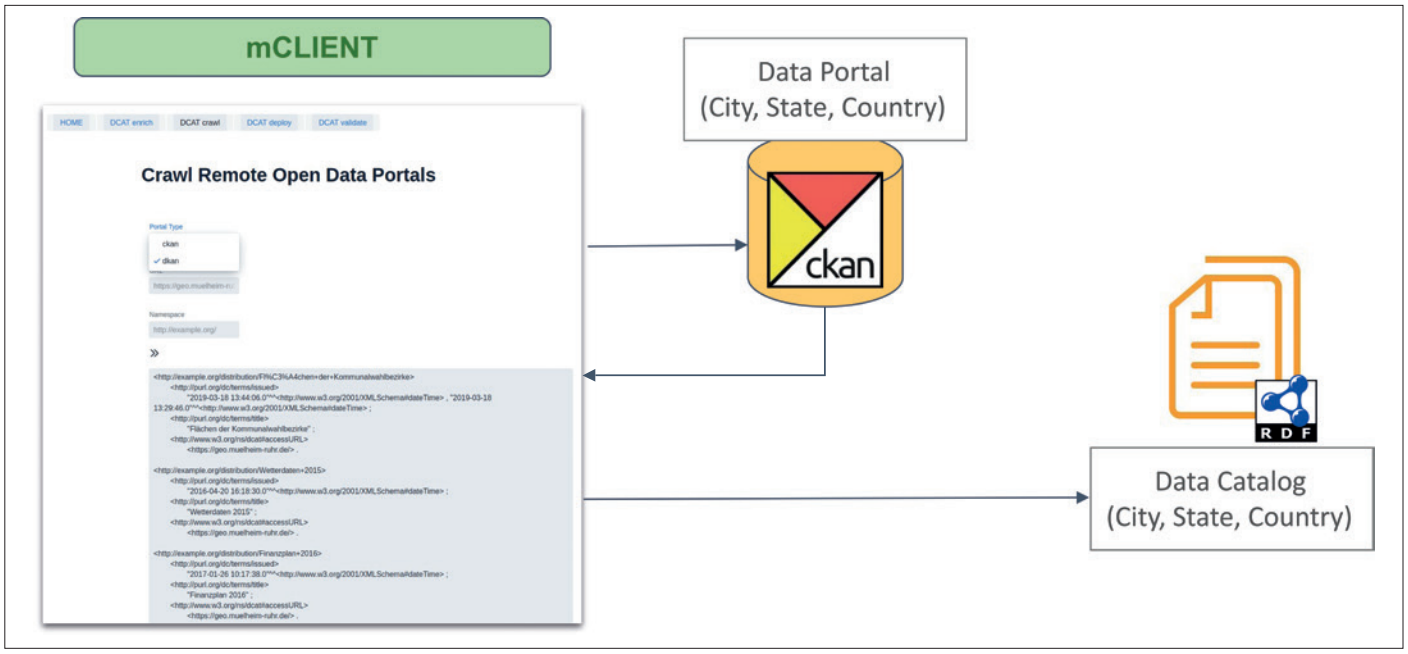


*Figure 4: mCLIENT crawling component*

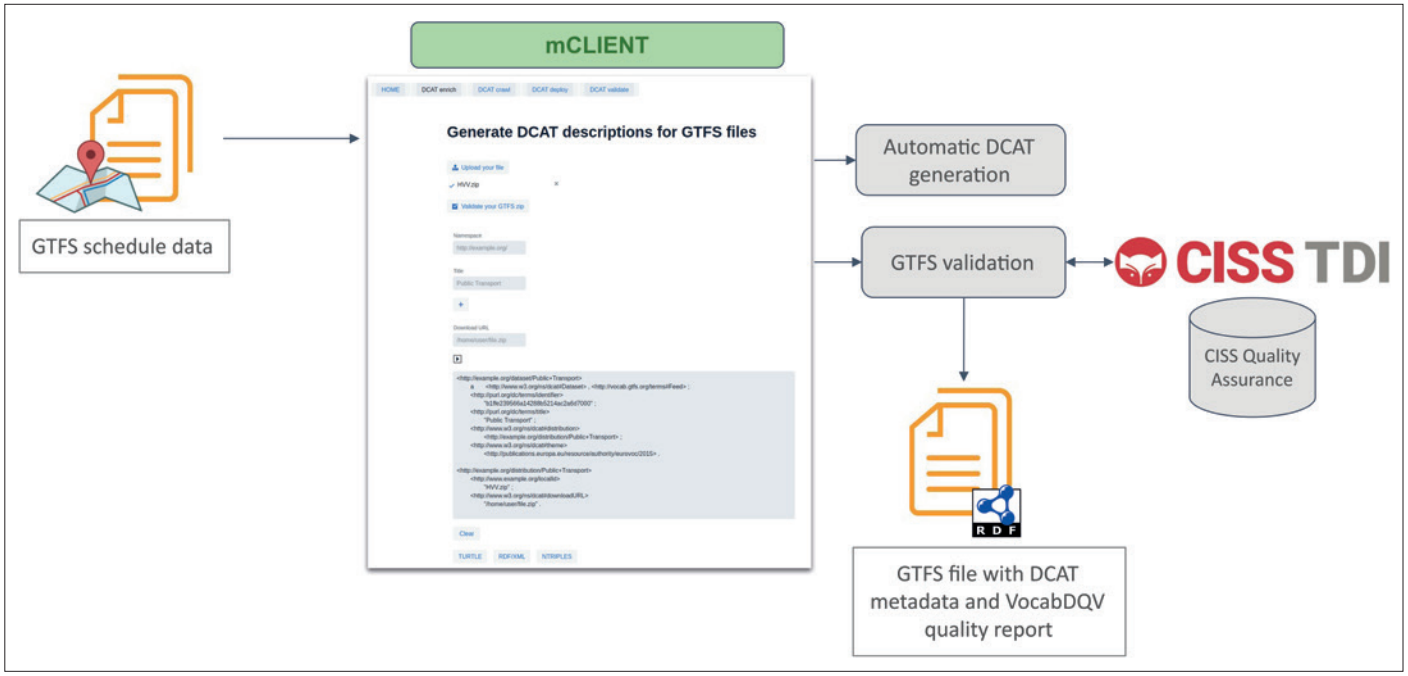


*Figure 5: mCLIENT data handling component (data quality & metadata generation)*

highlighting the need for automated verification of public transport data, we integrated a quality assurance process for GTFS data into the demonstrator.

We successfully tested the functionalities of the data handling component on the GTFS dataset of the public transport network of Hamburg (HVV)[6] and utilized it to automatically generate DCAT metadata. Figure 5 shows the generated metadata in the web interface of the data handling component.

An integration with the CISS QuAC ensures automatic quality checking is performed across various validation metrics from the Canonical GTFS Schedule Validator, which includes checks for empty files, duplicate columns, and overlapping stop times, as well as a geometric check for coordinate flips[7]. This is implemented as a web service that provides machine-readable reports and operates in a non-blocking manner. The process begins with registering a validation process on the server and generating a

[6] https://www.hvv.de/de/fahrplaene/abruf-fahrplaninfos/datenabruf

[7] https://gtfs-validator.mobilitydata.org/rules.html

process ID. The GTFS file is then uploaded to the web service, and the progress of the validation is monitored through queries. Finally, the results can be retrieved in the form of a VocabDQV RDF report[8]. The error report generated after validating the GTFS data via the mCLIENT helps data providers identify errors and contributes to improving the quality of mobility data in the long term. This, in turn, enhances trust in the provided data collections.

In a subsequent step of the data handling workflow, the generated metadata can be enriched with additional keywords. For this purpose, a portion of the metadata description (e. g., the title) is extracted and sent to the DBpedia Spotlight web service. The web service identifies relevant keywords from the DBpedia knowledge base, described in the form of Uniform Resource Identifiers (URIs). The returned keywords are subsequently added to the DCAT description. If the automatic annotation results in incorrect keywords, data providers can easily remove them from the DCAT snippet.

In addition to metadata enrichment, a check for the syntactic validity of the DCAT description can also be performed. For this purpose, the web service of the ITB DCAT-AP.de Validator is used. Figure 6 illustrates the appearance of such a validity check in the mCLIENT web interface.

For the HVV schedule data, the validator checked the completeness and syntactic correctness of the data, and also provided warnings and information regarding additional recommended or optional metadata fields.

### 4.3 DEPLOYMENT

The deployment component is used for publishing data collections on CKAN portals. Using the web interface of the mCLIENT demonstrator, datasets and their respective DCAT descriptions can be automatically published to Open Data portals with the push of a button. The data provider only needs the appropriate administrative rights on the respective data portal, which are granted by an API key. After successfully uploading a DCAT file along with the corresponding data collection, the publication process is initiated by clicking the "Deploy to CKAN" button. This triggers the mCLIENT to transfer the data collection, including the DCAT description, to the data portal. Figure 7 shows the web interface of the deployment component and the result of a successful publication of GTFS data from the public transport network for Central Germany (MDV) on a CKAN instance[9]. Currently, the deployment process only works for CKAN instances. However, integration with data portals using the DKAN interface will be possible with an extension of mCLIENT.

## 5 CONCLUSION AND OUTLOOK

Despite the potential benefits of Open Data, studies highlight significant challenges related to dataset and metadata quality, often leading to incomplete, outdated, or non-interoperable data being published. Tools like CKAN and DKAN offer some support through APIs and command-line interfaces to facilitate data publishing, but integration and automation gaps remain. Enhanced tool support for automated quality checks, metadata generation, and seamless deployment is crucial for improving the Open Data publishing workflow. Our qualitative survey among transport associations and Open Data portal operators from Germany, confirmed several challenges in the data publication process identified in the literature. The experts emphasized the need for automated workflows to improve quality and accessibility of open

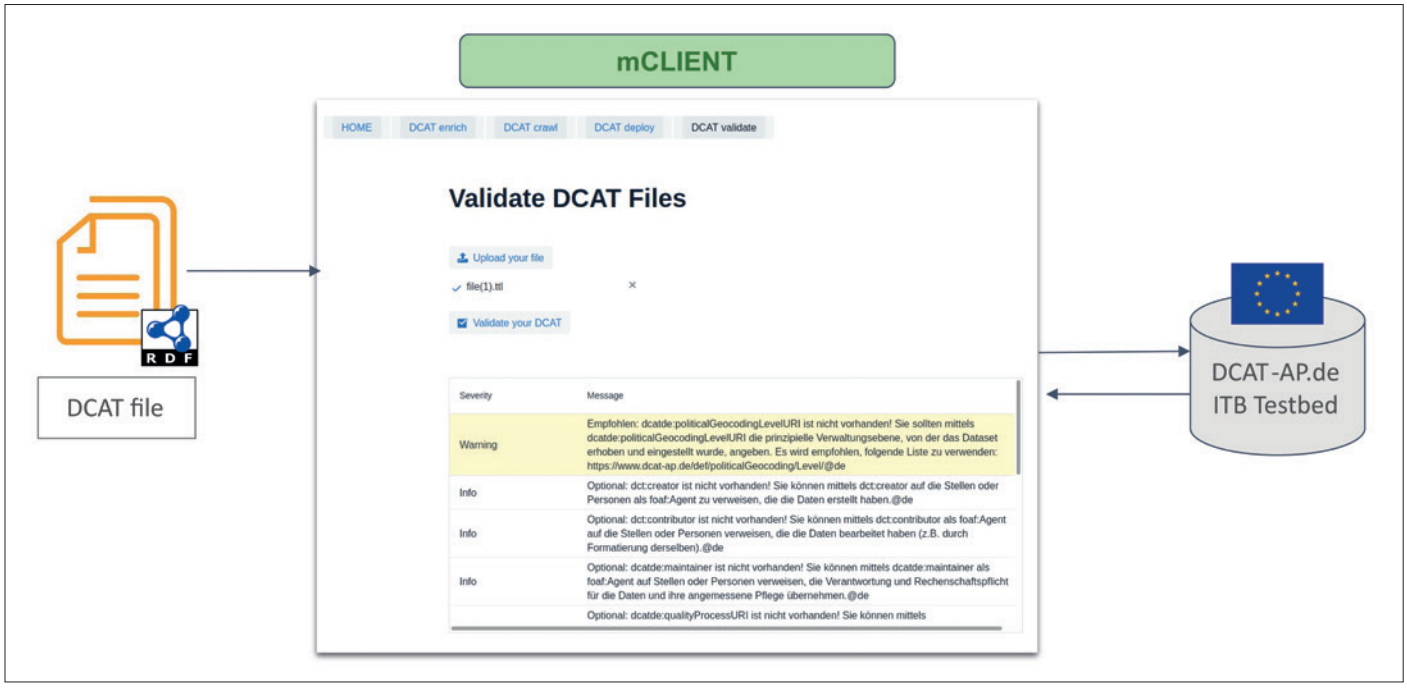


*Figure 6: mCLIENT data handling component – metadata validation*

[8] https://www.w3.org/TR/vocab-dqv

[9] https://www.govdata.de/daten/-/details/soll-fahrplandaten-mdv

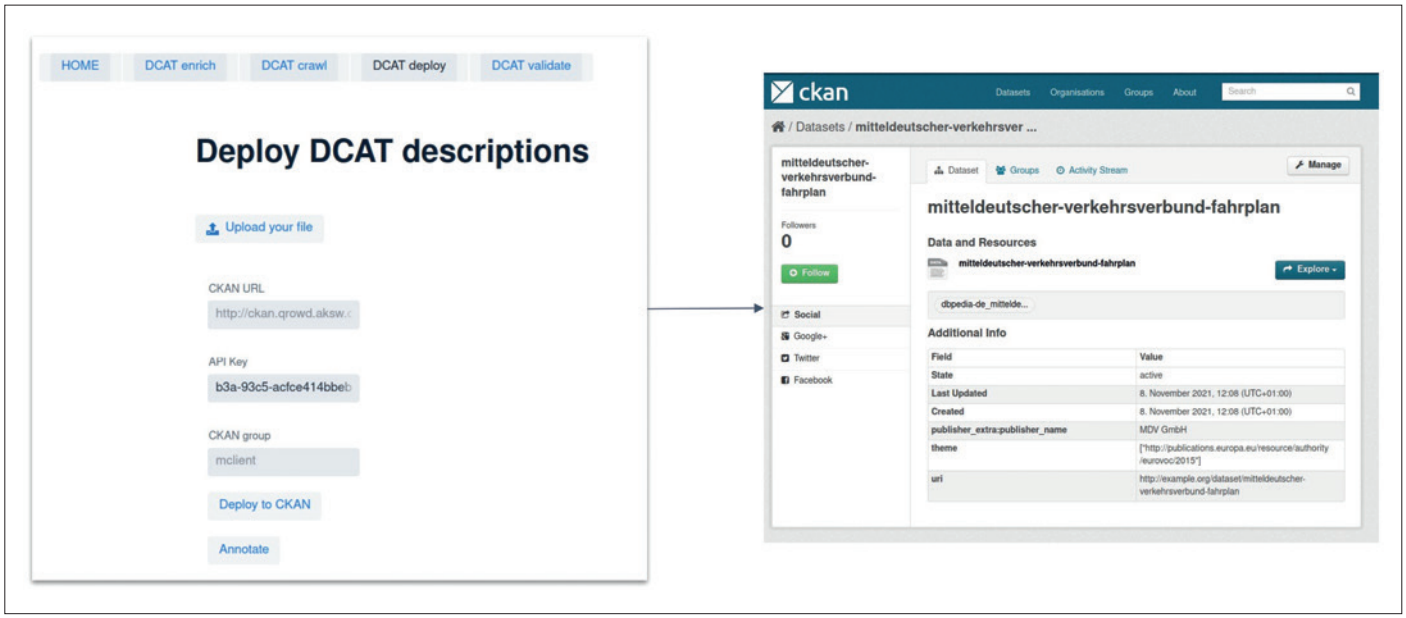


*Figure 7: mCLIENT deployment component*

datasets. The mCLIENT demonstrator addresses these challenges by effectively streamlining the data publication process. This automation not only improves data quality and trust but also significantly reduces the manual effort required for metadata generation and validation. The functionality of the system's components was successfully tested using sample GTFS data.

However, it is important to note that generating metadata, especially for text-heavy fields like titles or descriptions, still requires active participation from the data provider. If the corresponding texts are unclear or incomplete, these issues can propagate into the automatic annotation, potentially resulting in errors. To address this issue, data providers can use the mCLIENT software to remove incorrectly annotated keywords from the metadata description. A potential future measure to further enhance and enrich text-based metadata fields could involve integrating an API connection to large language models for supporting automatic text generation. To further enhance the annotation process and the interoperability of metadata, integrating additional web services is conceivable. These services could facilitate annotation using other knowledge graphs relevant to mobility data. For instance, GeoNames could be leveraged to incorporate spatial URI annotations into dataset descriptions. Additionally, the mCLIENT demonstrator could be extended to support file types beyond GTFS. For example, expanding its capabilities to include (real-time) traffic data, such as those adhering to the DATEX-II standard, would be beneficial. Such data could greatly benefit from automated error detection and correction.

Although the demonstrator is specifically designed for mobility data, our approach is versatile and can be readily adapted to other file types and application domains, given its generally applicable workflow.

## Acknowledgements

This work has been supported by the Federal Ministry of Transport and Digital Infrastructure (BMVI) under the grant number 19F2152A.

## References

Akyürek, H.; Scholl, C.; Stodden, R.; Siebenlist, T.; Mainka, A. (2018): Maturity and usability of Open Data in North Rhine-Westphalia. In: Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age, pp. 1–10.

Barry, E.; Bannister, F. (2014): Barriers to Open Data release: A view from the top. In: Information Policy 19 (1-2), pp. 129– 152.

Bill, R.; Lorenzen-Zabel, A.; Hinz, M. (2018): Offene Daten für Lehre und Forschung in raumbezogenen Studiengängen – OpenGeoEdu. In: gis.Science 1/2018, S. 32–44.

Conradie, P.; Choenni, S. (2014): On the barriers for local government releasing Open Data. In: Government Information Quarterly 31, pp. S10–S17.

European Commission (2023): Commission Implementing Regulation (EU) 2023/138.

Hinz, M.; Bill, R. (2018): Mapping the landscape of Open Geodata. In: Geospatial Technologies for All: Short Papers, Posters and Poster Abstracts of the 21st AGILE Conference on Geographic Information Science.

Hinz, M.; Bill, R. (2021): Exploring Open Data Portals for Geospatial Data Discovery Purposes. In: Kamilaris, A.; Wohlgemuth, V.; Karatzas, K.; Athanasiadis, I. N. (Eds.): Advances and New Trends in Environmental Informatics: Digital Twins for Sustainability. Springer International Publishing, Cham, Switzerland, pp. 147–162.

Janssen, M.; Charalabidis, Y.; Zuiderwijk, A. (2012): Benefits, adoption barriers and myths of Open Data and Open Government. In: Information Systems Management 29 (4), pp. 258–268.

McBride, K.; Toots, M.; Kalvet, T.; Krimmer, R. (2018): Open Government Data driven co-creation: Moving towards citizen-government collaboration. In: Electronic Government: 17th IFIP WG 8.5 International Conference, EGOV 2018, Proceedings, pp. 184–195.

Mendes, P. N.; Jakob, M.; García-Silva, A.; Bizer, C. (2011): DBpedia spotlight: Shedding light on the web of documents. In: Proceedings of the 7th International Conference on Semantic Systems, pp. 1–8.

Moyano, J. F. M.; López Beltrán, N. E.; Velandia Vega, J. A. (2017): Assessing data quality in Open Data: A case study. In: 2017 Congreso Internacional de Innovacion y Tendencias en Ingenieria (CONIITI). IEEE, Piscataway, NJ.

Neumaier, S.; Umbrich, J.; Polleres, A. (2016): Automated quality assessment of metadata across Open Data portals. In: Journal of Data and Information Quality (JDIQ) 8 (1), pp. 1–29.

Nikiforova, A. (2020): Open Data Quality Evaluation: A comparative analysis of Open Data in Latvia. In: arXiv preprint arXiv:2007.04697.

Wenige, L.; Stadler, C.; Bin, S.; Bühmann, L.; Junghanns, K.; Martin, M. (2020): Automatic Subject Indexing with Knowledge Graphs. In: LASCAR Workshop at the Extended Semantic Web Conference (ESWC 2020).

Wenige, L.; Stadler, C.; Martin, M.; Figura, R.; Sauter, R.; Frank, C. W. (2021): Open Data and the Status Quo – A Fine-Grained Evaluation Framework for Open Data Quality and an Analysis of Open Data Portals in Germany. In: arXiv preprint arXiv:2106.09590.

Wiedemann, M.; Bürger, T. (2021): Offene Verwaltungsdaten: Das Gemeinschaftsprojekt Musterdatenkatalog bietet erstmals einen Überblick. In: Kommunales Open Government: Grundlagen, Praxis, Perspektiven, p. 33.

Zhao, K.; Tarkoma, S.; Liu, S.; Vo, H. (2016): Urban human mobility data mining: An overview. In: 2016 IEEE International Conference on Big Data (Big Data), December 2016. IEEE, Piscataway, NJ., pp. 1911– 1920.

## IMPRESSUM // IMPRINT

gis.Science – Die Zeitschrift für Geoinformatik
ISSN 1869-9391

**Redaktion:**
Gerold Olbrich,
olbrich@vde-verlag.de,
Tel.: +49-69-840006-1121

**Hauptschriftleiter:**
Prof. Dr.-Ing. Ralf Bill,
ralf.bill@uni-rostock.de

**Editorial Board:**
Prof. Dr. Lars Bernard, TU Dresden; Prof. Dr. Thomas Brinkhoff, Jade Hochschule Oldenburg; Dr. Andreas Donaubauer, TU München; Prof. Dr. Max Egenhofer, University of Maine Orono; Prof. Dr. Manfred Ehlers, Universität Osnabrück; Prof. Dr. Klaus Greve, Universität Bonn; Prof. Dr. Martin Kada, Technische Universität Berlin; Dr. Stefan Lang, Universität Salzburg; Prof. Dr. Stephan Nebiker, Fachhochschule Nordwestschweiz; Prof. Dr. Pascal Neis, Hochschule Mainz; Prof. Dr. Josef Strobl, Universität Salzburg

**Internet:**
www.gisPoint.de

**Anzeigen:**
Tammy Rößler,
VDE VERLAG GMBH,
Tel.: +49-69-840006-1341,
roessler@vde-verlag.de

**Verlag:**
Wichmann Verlag im VDE VERLAG GMBH,
Bismarckstraße 33, 10625 Berlin,
Tel.: +49-30-348001-0,
Fax +49-30-348001-9088,
www.vde-verlag.de

**Geschäftsführung:**
Dr.-Ing. Stefan Schlegel

**Verlagsleiter Zeitschriften:**
Dipl.-Ing. Ronald Heinze

**Druck:**
Druck- und Verlagshaus Thiele & Schwarz GmbH, Kassel

**Anschrift für Zeitschriftenabonnements:**
WU.SOLUTIONS GmbH & Co. KG,
Große Hub 10, 63344 Eltville am Rhein,
Tel.: +49-6123-9238-234,
Fax +49-6123-9238-244,
vde-leserservice@vuservice.de

**Erscheinungsweise:**
10 x jährlich, davon 4 Ausgaben gis.Science, 6 Ausgaben gis.Business

**Jahresabonnement (4 Hefte):**
146,00 EUR zzgl. Versandkosten, Studenten/Auszubildende 61,50 EUR zzgl. Versandkosten, Mitglieder des Deutschen Dachverbands für Geoinformation e. V. (DDGI) erhalten das Abo im Rahmen ihrer Mitgliedschaft

**Bezugszeitraum:**
Ein Abonnement gilt für mindestens ein Jahr und verlängert sich jeweils um weitere 12 Monate, wenn es nicht bis spätestens 6 Wochen vor Ablauf des Bezugszeitraums gekündigt wurde. Reklamationen für nicht erhaltene Hefte können nur innerhalb von drei Monaten nach Erscheinen angenommen werden.



Die gis.Science ist seit 2004 in der internationalen Zitationsdatenbank Scopus gelistet.