Data Observer

Gotthard Meinel, Sujit Kumar Sikder* and Tobias Krueger

IOER Monitor: A Spatio-Temporal Research Data Infrastructure on Settlement and Open Space Development in Germany

Abstract: This paper gives a comprehensive introduction to the IOER Monitor – an open research data infrastructure (RDI) in Germany providing domain-specific multi-temporal geospatial datasets, services and visualizations for land use and land cover (LULC)-related development of settlements and open space and closely related topics. Its easy-to-use information system provides multi-scale data offers to form a discussion platform that supports spatial development assessment and evidence-based decision making. It contributes to public land-use change discourses by enhancing information offers that can be adopted by other multi-disciplinary data users - even from non-spatial domains. All data and services are freely available. IOER Monitor is committed to offering continuous services implementing FAIR principles (findable, accessible, interoperable and re-usable) and policy-relevant inputs for transformative spatial development.

1 Introduction

The land is one of the limited natural resources on our Earth. Only an efficient allocation and intensity of land use can ensure our livelihood supports and

*Corresponding author: Sujit Kumar Sikder,** Leibniz Institute of Ecological Urban and Regional Development, Weberplatz 1, 01217 Dresden, Germany, E-mail: s.sikder@ioer.de. https://orcid.org/0000-0002-0265-7394
Gotthard Meinel and Tobias Krueger, Leibniz Institute of Ecological Urban and Regional Development, Weberplatz 1, 01217 Dresden, Germany, E-mail: g.meinel@ioer.de (G. Meinel), t.krueger@ioer.de (T. Krueger). https://orcid.org/0000-0002-9201-7664 (G. Meinel). https://orcid.org/0000-0002-7085-8155 (T. Krueger)

necessary environmental protection (GLTN 2018). In particular, the physical development has to be carried out with careful considerations in case of new land consumptions.

The article aims to present a comprehensive introduction to an open research data infrastructure (RDI) which is providing domain-specific multi-temporal geospatial datasets, services and visualizations for Germany. The Monitor of Settlement and Open Space Development (IOER Monitor; https://www.ioer-monitor.de/en/) is a permanent RDI on the topic of land use/land cover which is hosted by the Leibniz Institute of Ecological Urban and Regional Development (IOER) in the city of Dresden, Germany. The IOER Monitor was established in 2010 as a scientific service of the IOER and has since developed over a decade into a comprehensive and sought-after research data infrastructure on the topic of land-use monitoring. It provides detailed spatial tangible information on land use in Germany, covering aspects like settlement structures, land take values, ecosystem services, spatial impacts of renewable energy fuels or landscape quality, just to mention a few. The data is provided for the whole terrestrial extent (i.e. land surface including inland water bodies and coastal areas) of Germany and is available for administrative units ranging from sub-municipal divisions to the federation and in most cases for INSPIRE-geographical grids from 100 to 10,000 m grid cells.

The subject-specific research data infrastructure presented in this article is fulfilling some of the pre-requisites of open data and e-Science principles, which was almost impossible without automatic workflow, systematic data management and integration of innovative information technologies. In recent time, "Open Science" or "Science 4.0" is becoming a basic requirement in science practice as well as in funding. Following such a trend, the innovation in digital technologies, communications and collaborative tools enable collecting, analysing, discovering and communicating essential research inputs by following a robust and plausible workflow. Today, the e-Science initiatives are trying to ensure participation of multidisciplinary or even transdisciplinary communities for tackling the global challenges together, given that the progress of science is key to innovation, growth and development pathway (Pardo Martínez and Poveda 2018). The emerging approaches of open science are often mentioned: (i) democratic, (ii) pragmatic, (iii) infrastructure, (iv) public and (v) measurable (Fecher and Friesike 2013). The key components of open science are open access, open data and open source (Jomier 2017). Pisani et al. (2016) suggested to think even beyond such limited open components rather highlighted many other potential basic requirements – e.g. equitable collaboration, supportive infrastructure, investment for the future data scientist, shared governance, interpretation and so on. In context, data providing infrastructures need to meet domain-specific community-driven principles like FAIR (findable, accessible, interoperable and re-usable), FAIR Data Maturity Model (RDA 2020) or CARE

(collective benefit, authority to control, responsibility and ethics) principles (GIDA 2019). Indeed, the national–global initiatives and platforms (including public-private partnerships) are communicating and establishing required guidelines, infrastructure and governances. Some examples are several German national research data infrastructures (NFDIs), which are currently under construction, the European Open Data Cloud (EOSC), GAIA-X and the Research Data Alliance (RDA).

2 Contemporary Challenges of Spatio-Temporal Monitoring

A long historical trend can be found of optimizing scarce resources like land in human history. However, there are evidence of unconscious land development, conversion and use for many reasons. In Germany, the coverage of settlement and transport-related land is following a rising trend even in regions of demographic shrinking. The statistical offices of the federal states and the federation annually provide basic statistics on land use and land-use change. However, this information is always given as aggregated numbers with the lowest granularity of singular municipalities. Though, much more information is often needed going deeper into the actual localities. The challenges of land monitoring are described in Meinel (2020) and solutions for this problem in Krüger et al. (2013). This information must be generated by systematic analysis and geo-processing of basic geospatial datasets in combination with various other data sources. For instance, the national climate protection plan of Germany aims to achieve the net-zero target for land consumption by 2050. A sustainable settlement development (i.e. low-to-zero land consumption, the greening of cities and avoidance of urban sprawl by prioritizing inner before outer development) requires more highly accurate as well as small-scale information on spatial features such as land use, buildings, urban greenery and transport infrastructure. The demand for accuracy of such data comprises aspects of geometrical accuracy, semantical correctness and topicality.

The need for small-scale information is diverse. This enables science and practice to optimize related measures and decisions, but it also increases efficiency to identify successes or mistakes. In particular, the challenges should be considered about the data – e.g. spatial and temporal resolution, objectivity, homogeneity and so on.

Potential data users are actors from the local, regional, national and even international level. The applications of the data are, for example, measuring coverages of different land-use classes on multiple scales, measuring land-take, quantifying landscape quality metrics, localizing hotspots of renewable energy

generation, estimating green volumes, building urban air pollution scenarios, estimating property values or geothermal potential calculations.

3 Research Data Infrastructure: IOER Monitor

3.1 Conceptual Framework of Settlement Open Space Monitoring

In general, the IOER Monitor follows the well-known principles of the DPSIR framework. Figure 1 shows a conceptual framework on land use monitoring that includes land as an environmental resource, spatial scale of settlement/open space and diverse factors of the social domain.
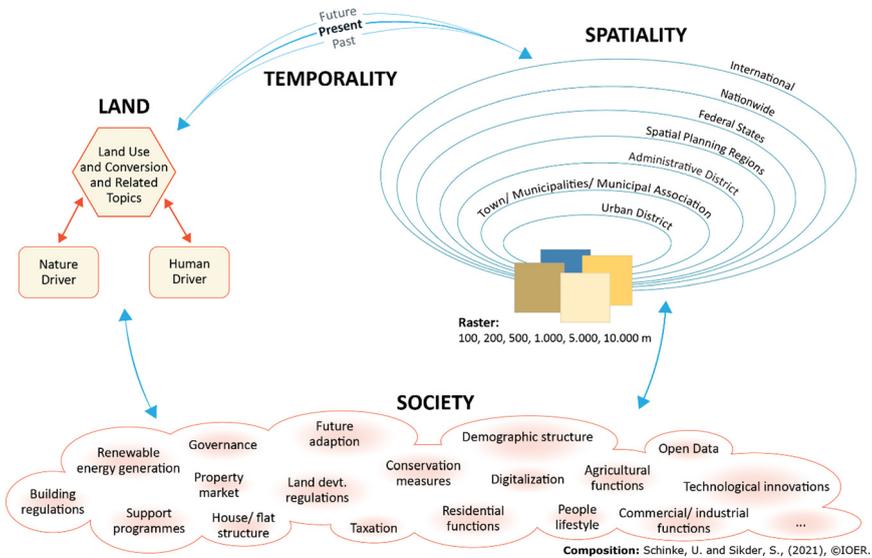
Figure 1: Conceptual framework of monitoring (source: Authors own).

As mentioned before, the geospatial scientific data infrastructure of the IOER Monitor offers a thematically wide range of datasets dealing with numerous aspects of land use, settlement structures and patterns. Moreover, it is the basic understanding of the authors of the IOER Monitor that datasets are not only offered for download but those indicators are presented interactively at multiple scales and prepared both graphically and in tabulated visualization. Hence, an easy-to-

use online platform has been developed to support scientists as well as decision-makers or the interested public with relevant information. Various usage needs from different user communities are met by various data presentation modes, such as maps and diagrams for presentation, or downloadable data tables for further individual investigation.

3.2 Thematic Indicators and Spatio-Temporal Resolution

This infrastructure provides comprehensive information on the status of land use and its changes through long-term time series in Germany at a high spatial resolution. The time series currently comprises 13 time periods beginning in 2000, then in 2006 and annually from 2008 onwards. The IOER Monitor is enabling an assessment of the sustainable land use development targets on all administrative levels including urban districts by means.

Currently, there are 95 indicators available, grouped into 10 categories such as settlement, transport, open space, sustainability, building/material stock, urban sprawl, landscape quality, ecosystem services, renewable energy and risk.

Most indicators describe the proportions of several land-use categories of different reference areas. Land use categories are arranged in a pre-defined land use classification scheme, which is hierarchically dividing the land surface into several unique land use classes.

However, besides these area-proportional indicators (i.e. built-up area percentage), there are also more complex indicators such as population-related numbers (settlement density), indicators for the quantification of landscape quality (e.g. hemeroby index, landscape fragmentation and urban sprawl), eco-system service supply (urban green space per inhabitant and potential urban green accessibility) or indicators for measuring the dynamics of landscape change (i.e. land consumption rate of settlement and transportation per day). Further indicators are targeting building stocks or development in flood-risk prone areas, for instance.

3.3 Data Mining, Geoprocessing, Harmonization and Curation

The preparation of end-user information and dataset has to deal with intensive computation due to the related complexity of high volume of data, the various data providers and lack of harmonization.

The data sources for land-use related indicators are annual editions of the Basic Digital Landscape Model of the Authoritative topographic–cartographic

information system (ATKIS-Basic DLM, base scale ~1:25,000). The basic DLM is the harmonized digital landscape model for all federal states providing detailed geo-topographic area-covering and gap-less mesh-wise spatial description of Germany (e.g. Krueger et al. 2013; Schorcht et al. 2016). Besides polygons representing areal land-surface objects, it also features linearly modelled transport infrastructure (esp. the road and railway networks).

Further important input data sets for the IOER Monitor include 2D- and 3D-building models (HU-DE, LoD1-DE) featuring building polygons and blocks derived from cadastral data, the German land cover model (Landbedeckungsmodell Deutschland LBM-DE), which is provided by the Federal Agency for Cartography and Geodesy (Bundesamt für Kartographie und Geodäsie BKG) in a three-year cycle and is the official German contribution to the European CORINE Land-Cover products.

For retrospective analyses, approaches to automatically extracting information from digitized topographic maps 1:25,000 (Topographische Karte TK25) and its historical predecessors have been developed (Herold 2018). Further geo-data which is part of the geo-processing are, for instance, protected areas for nature conservation, flood-risk areas, the High-Resolution Layer Imperviousness Degree of the Copernicus Land Monitoring Service as well as official statistical information (e.g. population, traffic, economy and finance). Buildings are the important information carriers of settlements. The extraction and classification of buildings from topographic maps are described in Hecht et al. (2015).

Each of the datasets requires specific pre-processing steps for data harmonization. Most importantly, the geo-data has to be transformed into a consistent geographical projection. Since most indicators are area-related the Lambert azimuthal equal-area projection is used for data analysis.

Concerning land use, relevant objects from the ATKIS Basic DLM object group Real Usage are collected and assigned to one of the 35 categories from the land use scheme (Krueger et al. 2013). Linear objects representing features with significant land cover proportions (roads, railway infrastructure and running waters) have to be transformed to polygons by buffering. Overlapping conflicts of different polygons are solved by a pre-defined prioritization scheme.

For most of the indicators, an automated workflow has been set up for the geoprocessing tasks at the federal state level. In the end, the partial geometries of the federal states are merged as an overall data set and used for further computations at the different spatial scales (sub-municipal districts to the whole country, raster cells). The cutting edge innovative geo-spatial technologies are rapidly adopted from the powerful analysis tools of ESRI ArcGIS, FME as well as open-source software packages (QGIS). The developed solutions are suitable for high-performance geo-processing of very large data sets through a combination of tiling and parallel processing of analytical steps into individual sequences and so on.

3.4 Description of Data Quality and Sensitivity

According to international guidelines, RDI repositories should provide enough information about the data and related uncertainties for the users to assess the quality of the data (CoreTrustSeal-SCB 2019). IOER Monitor Indicators are regularly calculated according to the cyclical availability of their corresponding input datasets. For instance, land-use related indicators are calculated annually since the ATKIS Basic DLM is provided every year (in fact there is a quarterly delivery of the data sets from the federal states' mapping agencies to the BKG). Landscape quality indicators such as hemeroby index and near-natural area percentage are calculated in a three-year interval corresponding to the input data LBM-DE (BKG 2021). Uncertainty in the application of LBM-DE in monitoring is explained in Meinel and Reiter (2019). For other indicators a perennial monitoring interval is chosen for thematic specific reasons, e.g. landscape fragmentation is calculated every four years. The geodata sets on which the indicators are calculated are intersected with the reference geometries after their completion to calculate the indicator values for the different scale levels. Most indicators are calculated both for administrative units (municipalities, districts and federal states) and at the grid level (cell sizes 100–10,000 m). In order to achieve multi-temporal comparability, earlier periods are calculated to the latest available administrative-territorial status. This compensates for changes in municipal affiliation that would otherwise lead to difficulties in the comparative analysis.

IOER Monitor provides comprehensive documentations and even case-by-case support to evaluate the completeness of the dataset and emerging sensitivity for any potential user. Detailed documentation of the methods, data sources, assumptions, related literature citations are making available and regularly updated on the monitor homepage for each indicator.

3.5 Analytics, Geo-Services and Metadata Catalog

(i) Interactive Geovisualization, Statistics and Predictions:

An interactive, internet-based geo-viewer enables the user to obtain digital cartographic visualization, statistics and analyses at a desired spatial scale such as federal state, spatial development regions (Raumordnungsregionen) as they are defined by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (Bundesinstitut für Bau-, Stadt- und Raumforschung BBSR), district, municipality, urban district and INSPIRE-compliant raster grids (cell sizes of 100, 200, 500, 1000, 5000 and 10,000 m). The selected indicator values are displayed as

interactive maps, tables and diagrams. Comparative functionalities are enabled on the selected spatial scale, for example, an evaluation of how own city or region performs about a selected indicator. Points to be noted, a potential user does not require to log in for access to IOER monitor data, visualization and will not be tracked; and data and services are available free of cost. In addition, for each year a land-use map is calculated in 2.5 m raster resolution, which allows small-scale change detection when comparing different periods.

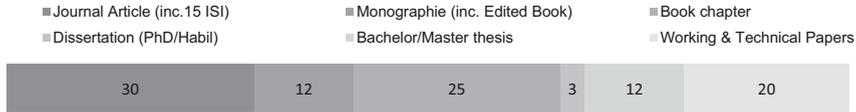(ii) Standard Open Geo-Spatial Service and RESTfull-API:

An advanced user should enjoy the display capability of the maps in the GIS environment as Web Map Service (WMS) geoservice; but can also be imported directly via geo-services such as Web Coverage Service (WCS) or Web Feature Service (WFS). Therefore, the user can conduct the required analysis in combination with other datasets and even can be directly integrated with their own web services. Since January 2021, in order to fulfil the requirement of the research data infrastructure certification principle a registration system has been established for providing a unique API-Key to users of geo-services (only for WFS and WCS) – where an easy and straight forward contract should be agreed by the user; however, all other services are open without registration.

(iii) Standard Metadata for Automated Harvesting:

INSPIRE-conform standard Metadata services are made available and updated regularly for all monitor data services. This structured metadata is compatible with automatic harvesting and linking by national and regional metadata portals such as Geoportal.de (www.geoportal.de), Geoportal Sachsenatlas (geoportal.sachsen. de) and mCloud (www.mcloud.de).

4 IOER Monitor for Science and Practice

Since 2010, IOER monitor data and services are already used and integrated for a significant number of scientific publications, applied research projects, research data infrastructures and education platforms. Multi-disciplinary researchers, professionals and data journalists are using data services and even seek support for special thematic research questions. Total of 117 scientific contributions (does not count externals) has been published with direct use of monitor data since 2010 (Figure 2).

■ Journal Article (inc.15 ISI)   ■ Monographie (inc. Edited Book)   ■ Book chapter
■ Dissertation (PhD/Habil)   ■ Bachelor/Master thesis   ■ Working & Technical Papers

30	12	25	3	12	20

Figure 2: Scientific publications records with IOER monitor data (till 2019).

(i) Contribution to long-term scientific infrastructures:

Some recent examples can be mentioned as (1) DFG funded long term scientific data infrastructure project Social and Spatial Research Data Infrastructure (Sozial-Raumwissenschaftliche Forschungsdateninfrastruktur SoRa) that is linking high dimensional data infrastructure for social and spatial science (Bensmann et al. 2020). In the context of the presented infrastructure for geospatial linking, IOER Monitor geoservices were added which can be used to link interviewed persons from social science surveys to their environment with regard to land use (e.g. share of settlement and traffic area, building density, imperviousness degree of soil) and accessibility of the public infrastructure (e.g. public transport stops and green areas). Socio-spatial requests are answered by accessing IOER Monitor raster type data and retrieve the relevant indicator values from the corresponding 100×100 m raster grid or are calculated directly (e.g. distance to the nearest public transport stop) by means of a WPS service using on-the-fly calculations based on the geodata.

   (2) The OpenGeoEdu project, funded by the Federal Ministry of Transport and Digital Infrastructure (BMVI) through the mFund research initiative, adopted IOER Monitor data, services and methodology for developing a best practice case study of land use monitoring (Bill et al. 2020; Sikder et al. 2019). The OpenGeoEdu project aimed to promote the use of open geodata for spatial science-oriented study courses by offering diverse e-learning modules for flexible adaptation in current academic and professional education.

(ii) Adoption in collaborative research projects:

IOER Monitor data and services are frequently adopted within several multidisciplinary collaborative projects; some examples where IOER research area "Monitoring" participated: urban green – meinGrün (Brzoska et al. 2020), urban air quality – SAUBER (Petry et al. 2020), diversity of insects in nature protected areas – DINA, building-integrated photovoltaic Standard-BIPV and by the Center for Scalable Data Analytics and Artificial Intelligence- ScaDS.AI.

   Additionally, IOER Monitor data and services often adopted in external research projects; however, specific details are hard to track due to open data policy and no obligations for permissions.

(iii) Adoption by planning practitioners and regional authority:

The data/services of the IOER Monitor are used in development policy formulation, public administration, urban and spatial planning, economic sector, science/education and even private individual interests. For example, a comparison of settlement development of urban regions in Germany (Siedentop and Meinel 2020), a monitoring of ecosystem services (Grunewald et al. 2017) and mobility infrastructure planning with land-use indicators (Sikder et al. 2018). The IOER Monitor team annually hosts the Dresden Land Use Symposium (Dresdner Flächennutzungssymposium – DFNS) to provide a discussion and exchange platform for practice-oriented thematic knowledge. Accompanying the annual symposium, a book series with selected contributions on current political discussions, practical solutions and technical innovations in the field of spatial data science is published: http://www.IOERMonitor.de/tagungen.

(iv) Recognition from data community:

Since 2020, IOER Monitor is a certified RDI by German Data Forum (RatSWD). RatSWD is an advisory panel to the federal and state governments on research data infrastructure in Germany. IOER is a registered research data repository on re3data.org to be found under the following DOI: https://doi.org/10.17616/R3QF5P

5 Conclusion

The IOER Monitor provides an easy-to-use information system for spatial planning, environmental administration and spatial science. The information offered help for political actors to assess developments and make evidence-based decisions, for example to contributes to the evaluation of the development of SDG indicators. It can also contribute to a further improvement in the quality of geo-topographic databases; therefore, the enhanced information offers can be adopted by other multidisciplinary data users, even non-spatial domains. All data and services are freely available under DE-CC-BY-2.0. Since the inauguration of IOER Monitor, it is contributing to the commitment of domain-specific spatio-temporal research data on settlement and open space development, which already fulfil many fundamental principles of open science and practice. In future, IOER Monitor will provide enhanced competence services as a permanent research data infrastructure (RDI) by implementing a higher level of FAIRness principles, requirements of designated community and certifications.

5.1 Important Information

Contact person für FDZ IOER-Monitor: Dr. Gotthard Meinel, Leibniz Institute of Ecological Urban and Regional Development, Weberplatz 1, 01217 Dresden, Phone +49 (0)351 46 79 0, E-mail: monitor@ioer.de

Leibniz Institute of ecological urban and regional development (IOER):
https://www.ioer.de/1/home
Data overview of IOER Monitor: https://www.ioer-monitor.de/en
Dresdner Flächennutzungssymposium (Annual Dresden Land Use Symposium): http://12dfns.ioer.info
International Landuse Symposium: http://ilus2019.ioer.info

References

Bensmann, F., Heling, L., Jünger, S., Mucha, L., Acosta, M., Goebel, J., Meinel, G., Sikder, S., Sure-Vetter, Y., and Zapilko, B. (2020). An infrastructure for spatial linking of survey data. Data Sci. J. 19: 27.

Bill, R., Lorenzen-Zabel, A., Hinz, M., Kalcher, J., Pfeiffer, A., Brosowski, A., Aberle, H., Hovenbitzer, M., Meinel, G., Sikder, S., et al. (2020). OpenGeoEdu – a massive open online course on using open geodata. ISPRS Ann. Photogrammetry Remote Sens. Spatial Inf. Sci. 5: 31–38.

BKG. (2021). Landbedeckungsmodell für Deutschland (LBM-DE), Bundesamt für Kartographie und Geodäsie, Available at: <https://www.bkg.bund.de/DE/Ueber-das-BKG/Geoinformation/Fernerkundung/Landbedeckungsmodell/landbedeckungsmodell.html>.

Brzoska, P., Hecht, R., and Stanley, C. (2020). Stadtgrün mit einer App (neu) entdecken. Transform. Cities 4: 86–90.

CoreTrustSeal – SCB. (2019). CoreTrustSeal trustworthy data repositories requirements 2020–2022 (version v02.00-2020-2022). CoreTrustSeal Standards and Certification Board, Available at: <https://zenodo.org/record/3638211#.YFNb-twxnIU>.

Fecher, B., and Friesike, S. (2013). Open science: one term, five schools of though. In: Friesike, S., and Bartling, S. (Eds.), *Opening science*. Springer, Cham, pp. 17–47.

GIDA. (2019). CARE principles for indigenous data governance. Research data Alliance International Indigenous Data Sovereignty Interest Group. The Global Indigenous Data Alliance (GIDA), Available at: <https://www.gida-global.org/care>.

GLTN. (2018). Land management and planning. Global Land Tool Network (GLTN), Available at: <https://gltn.net/home/land-management-and-planning/#land-use-planning>.

Grunewald, K., Syrbe, R.-U., Walz, U., Richter, B., Meinel, G., Herold, H., and Marzelli, S. (2017). Germany's ecosystem services – state of the indicator development for a nationwide assessment and monitoring. One Ecosyst. 2: e14021.

Hecht, R., Meinel, M., and Buchroithner, M. (2015). Automatic identification of building types based on topographic databases – a comparison of different data sources. Int. J. Cartogr. 1: 18–31.

Herold, H. (2018). *Geoinformation from the past: computational retrieval and retrospective monitoring of historical land use*. Springer Spektrum, Wiesbaden, Germany, pp. 1–144.

Jomier, J. (2017). Open science–towards reproducible research. Inf. Serv. Use 37: 361–367.

Krüger, T., Meinel, G., and Schumacher, U. (2013). Land-use monitoring by topographic data analysis. Cartogr. Geogr. Inf. Sci. 40: 220–228.

Meinel, G. (2020). Herausforderung Flächenmonitoring: Datenquellen für ein Flächeninformationssystem und was sie leisten können. Stadtforschung Stat. 33: 107–114.

Meinel, G. and Reiter, D. (2019), Nutzung des Landbedeckungsmodells LBM-DE für das Flächenmonitoring – Bewertung und Ergebnisse. In: Meinel, G., Schumacher, U., Behnisch, M., and Krüger, T. (Hrsg.). *Flächennutzungsmonitoring XI : Flächenmanagement – Bodenversiegelung – Stadtgrün*, 2019 (IÖR-Schriften; 77). Rhombos-Verlag, Berlin.

Pardo Martínez, C. and Poveda, A. (2018). Knowledge and perceptions of open science among researchers — a case study for Colombia. Information 9: 292.

Petry, L., Herold, H., Meinel, G., Meiers, T., Müller, I., Kalusche, E., Erbertseder, T., Taubenböck, H., Zaunseder, E., Srinivasan, V., et al. (2020). Air quality monitoring and data management in Germany – status quo and suggestions for improvement. In: *ISPRS TC IV – 5th international conference on smart data and smart cities,* Nice, France, 2020.

Pisani, E. (2016). Beyond open data: realising the health benefits of sharing data. BMJ 355: i5295.

RDA. (2020). FAIR Data Maturity Model. Specification and Guidelines (Version 1.0). Contributed by FAIR data maturity model working group. Global Research Data Alliance (RDA), Available at: <https://doi.org/10.15497/rda00050>.

Schorcht, M., Krüger, T., and Meinel, G. (2016). Measuring land take: usability of national topographic databases as input for land use change analysis: a case study from Germany. ISPRS Int. J. Geo-Inf. 5: 134.

Siedentop, S., Meinel, G., and Pehlke, D. (2020). Flächensparende Siedlungs-entwicklung in deutschen Stadtregionen – eine Bilanz. ILS Trends 03/20: 1–8.

Sikder, S., Herold, H., and Meinel, G. (2018). Halten Sie die Augen nach einer Ladestation offen! Die E-Ladestations-Infrastruktur in Deutschland: eine räumliche Analyse. AGIT – J. für Angew. Geoinformatik 4: 218–227.

Sikder, S., Herold, H., Meinel, G., Lorenzen-Zabel, A., and Bill, R. (2019). Blessings of open data and technology: e-learning examples on land use monitoring and e-mobility. In: Getzinger, G. and Jahrbacher, M. (Eds.). *Proceeding of critical issues in science, technology and society studies*. Verlag der Technischen Universität Graz, Austria.