

Multi-Sensor-Systeme

Forschungsdatenmanagement

Wintersemester 2022/2023

Dr.-Ing. Sören Vogel

„Forschungsdaten sind eine wesentliche Grundlage für das wissenschaftliche Arbeiten. Die Vielfalt solcher Daten entspricht der Vielfalt unterschiedlicher wissenschaftlicher Disziplinen, Erkenntnisinteressen und Forschungsverfahren. Zu Forschungsdaten zählen u.a. Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Surveydaten, Objekte aus Sammlungen oder Proben, die in der wissenschaftlichen Arbeit entstehen, entwickelt oder ausgewertet werden. Methodische Testverfahren, wie Fragebögen, Software und Simulationen können ebenfalls zentrale Ergebnisse wissenschaftlicher Forschung darstellen und sollten daher ebenfalls unter den Begriff Forschungsdaten gefasst werden. **Die langfristige Sicherung und Bereitstellung der Forschungsdaten leistet einen Beitrag zur Nachvollziehbarkeit und Qualität der wissenschaftlichen Arbeit und eröffnet wichtige Anschlussmöglichkeiten für die weitere Forschung.“**

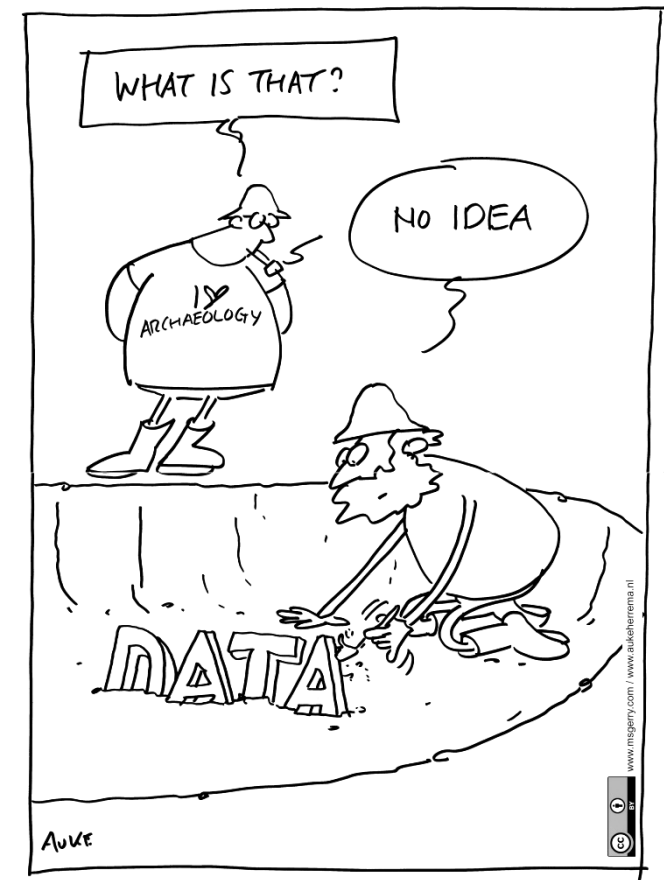
Deutsche Forschungsgemeinschaft (DFG) – 2020

- Aktivitäten und Informationen rund um die Forschungsdaten
- Bewusster und systematischer Umgang mit Daten von der ersten Idee bis zum Projektabschluss und ggf. darüber hinaus → Nachnutzbarkeit
- Planen, sammeln/generieren, verwalten, strukturieren, aufbereiten, dokumentieren, auswählen, analysieren, archivieren, publizieren von Forschungsdaten
- Zum Teil Vorgaben, Richtlinien, Gesetze
 - Datenschutzgesetze (DSGVO), Urheber- und Patentrechte, Leitlinien / Kodex zur Sicherung guter wissenschaftlicher Praxis der DFG, Richtlinien der LUH



[by Heinz-Vale]

- Strukturierten Überblick er- / behalten
- Zusammenarbeit erleichtern
- Qualität wahren
- Formale Anforderungen erfüllen
- Nachvollziehbarkeit / Auffindbarkeit gewährleisten
- Veröffentlichung von Forschungsdaten für
 - wissenschaftlichen Austausch
 - Förderfähigkeit durch Drittmittelgeber
 - Reputation und öffentliche Zugänglichkeit
 - Transparenz und Reproduzierbarkeit
- Langfristige Sichtbarkeit, Sicherheit und Verfügbarkeit von Forschungsergebnissen unter der Einhaltung datenschutzrechtlicher Vorgaben



DATA FOR FUTURE GENERATIONS

[by Auke Herrema]

FIAR-Prinzip:

- **Findable** (auffindbar)
 - Verwendung sinnvoller Metadaten und globaler Digital Object Identifiers (DOI)
- **Accessible** (zugänglich)
 - Open access und Nutzung gebräuchliche Protokolle / Formate
- **Interoperable** (interoperable)
 - Einhaltung und Anwendung fachspezifischer Standards / Sprachen
- **Re-useable** (wiederverwendbar / nachnutzbar)
 - Hilfreiche und nachvollziehbare Dokumentation mit genauen und relevanten Attributen

Für die Beschreibung und Dokumentation von Daten und deren Verarbeitung.
Strukturierte Informationen, um Daten zu verstehen und (nach-)nutzen zu können:

- Was ist das? → Kontext / Beschreibung
- Von wem stammt das? → Autor
- Woher kommt das ursprünglich? → Quelle, Speicherort
- Kann/Darf ich das (nach-)nutzen? → Lizenz
- Von wann ist das? → Datum der Erstellung / Publikation
- Kann ich das öffnen / lesen / bearbeiten? → Dateiformat

Realisierungen: readme.txt, strukturierte XML-Datei, File-Header, etc.

Beispiele für Metadaten: *Titel, Ersteller, Datum, Beschreibung, Identifikator, Beteiligte, Erfassungsbereich, Format, Sprache, Herausgeber, Beziehung, Rechte, Quelle, Thema, Typ, Größe, Version, (Geo)Lokalisierung, ...*

Strukturiertes, einfaches und schlankes Dokument mit detaillierten Informationen zum Umgang mit Daten. Kann und sollte im Projektverlauf immer wieder ergänzt und angepasst werden → kein statisches Dokument

Mögliche Themen:

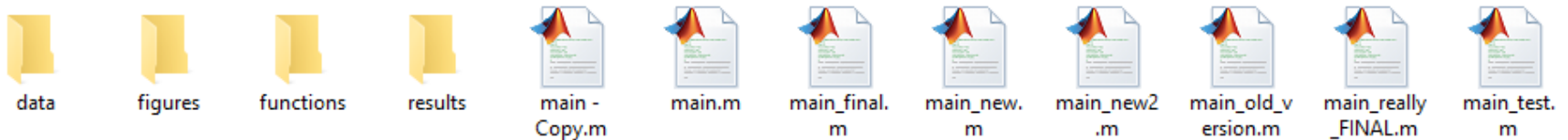
- Speicherung
 - Speicherkapazität, Bandbreite, Backup, Administrative Angaben, methodische Grundlagen, Verantwortlichkeiten
- Strukturierte Dateiablage und Benennung
 - Datum, aussagekräftiger Titel, Bearbeiter, Version
- Welche Daten langfristig aufbewahren?
 - Abhängigkeit u.a. vom Speicherplatzbedarf, der Qualität, dem Aufwand für eine etwaige Reproduktion sowie der Einzigartigkeit

Beispielhafte Inhalte (abhängig vom Projekt):

- Überblick mit Projektbeschreibung
- Datenbestand
 - Verwendete/Erfasste Daten, Größe, Formate, methodische Grundlagen
- Organisation und Sicherung der Daten
 - Benennung, Speicher-Plan, Backup-Plan, Archivierung
- Übersicht zur Erfassung/Prozessierung der Daten
 - Software, Hardware, Sensoren
- Veröffentlichen und Bereitstellen der Daten
 - Ggf. datenschutzrechtliche Aspekte

Nutzung von geeigneten Diensten:

- Git, Wiki, (Onlyoffice, Seafile) für geteilten Zugriff und Backup



Sicherlich eine Möglichkeit, aber:

- Was ist die aktuellste Version?
- Was sind die Unterschiede zwischen den einzelnen Versionen?
- Letzte lauffähige Version noch vorhanden?

→ Verwendung einer Versionsverwaltung

- Automatisches Backup
- Historie und Änderungsverlauf
- Vereinfachte Vereinigung von unterschiedlichen Zwischenständen
- Hauptsächlich geeignet für Textdateien (Code, Dokumentation, Konfigurationsdateien)

Was ist **Git**?

- Verteilte Versionsverwaltung (open source)
- Häufig für Quellcodeentwicklung genutzt
- Grundsätzlich für Änderungsverfolgung beliebiger Dateien geeignet



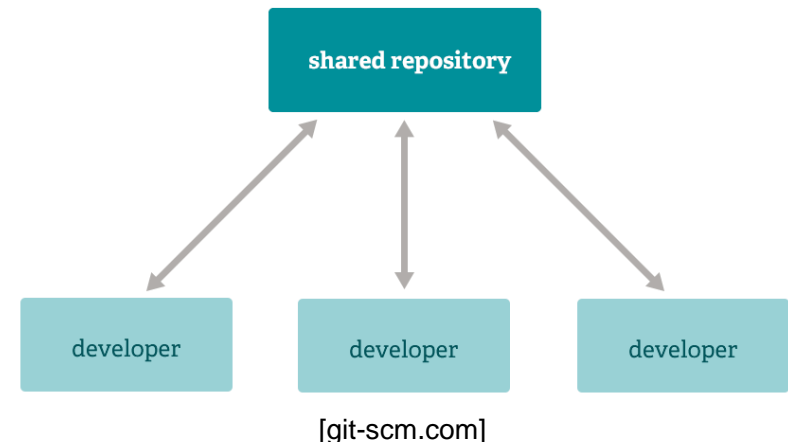
[git-scm.com]

Trivia

- Schlüsselfigur bei der Entwicklung von Git ist Informatiker und Softwareentwickler Linus Torvalds, Erfinder des Linux Kernels
- „git“ englisch (Brit.) umgangssprachlich für „Depp / Blödmann“
- Ursprüngliches Release im April 2005, aktuelle Versionsnummer ist 2.39.0 (vom 12.12.2022)

Wesentliche Eigenschaften:

- Dezentraler Aufbau
 - Jeder Nutzer arbeitet in lokaler Kopie („clone“)
 - Kopie beinhaltet gesamtes Repositorium inklusive Versionshistorie
- Nicht-lineare Entwicklung
 - Simple Abspaltung (engl. Branch) und
 - Verschmelzung (engl. Merge) einzelner unabhängiger Entwicklungszweige / -projekte
- Verschlüsselte Versionshistorie
- Betriebssystemübergreifend



Grafische Nutzeroberflächen

- Grafische Darstellung der Versions- und Commit-Historie
- Integrierter Code-Editor
- Unterstützung von Drag & Drop
- Einbindung von gängigen, webbasierten Git-Hosts



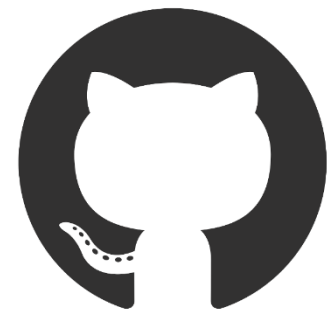
axosoft
GitKraken

[gitkraken.com/store#logos]



Git Extensions

[gitextensions.github.io]



GitHub

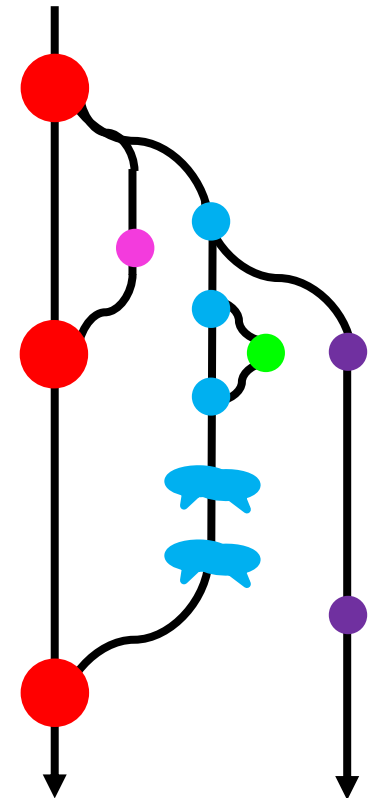
[github.com]

Beispielhafter Workflow:

- Branch anlegen, um Änderungen zu implementieren und zu testen
- Daten hinzufügen, bearbeiten oder löschen
- Veränderungen mit aussagekräftiger Beschreibung committen und damit Arbeitsprozess dokumentieren
- Veränderungen durch weitere Teammitglieder prüfen lassen und ggf. Anpassungen vornehmen
- Änderungen testen
- Branch mit 'master' zusammenlegen und kombinieren

Hinweise:

- Im Allgemeinen aussagekräftige Bezeichnungen verwenden sowie prägnante und informative Kommentare verfassen
- In kleinen Schritten arbeiten
 - Nicht mehrere Änderungen, die nicht in Zusammenhang stehen, gleichzeitig committen
 - Stattdessen kleinere, logisch in sich geschlossene Commits

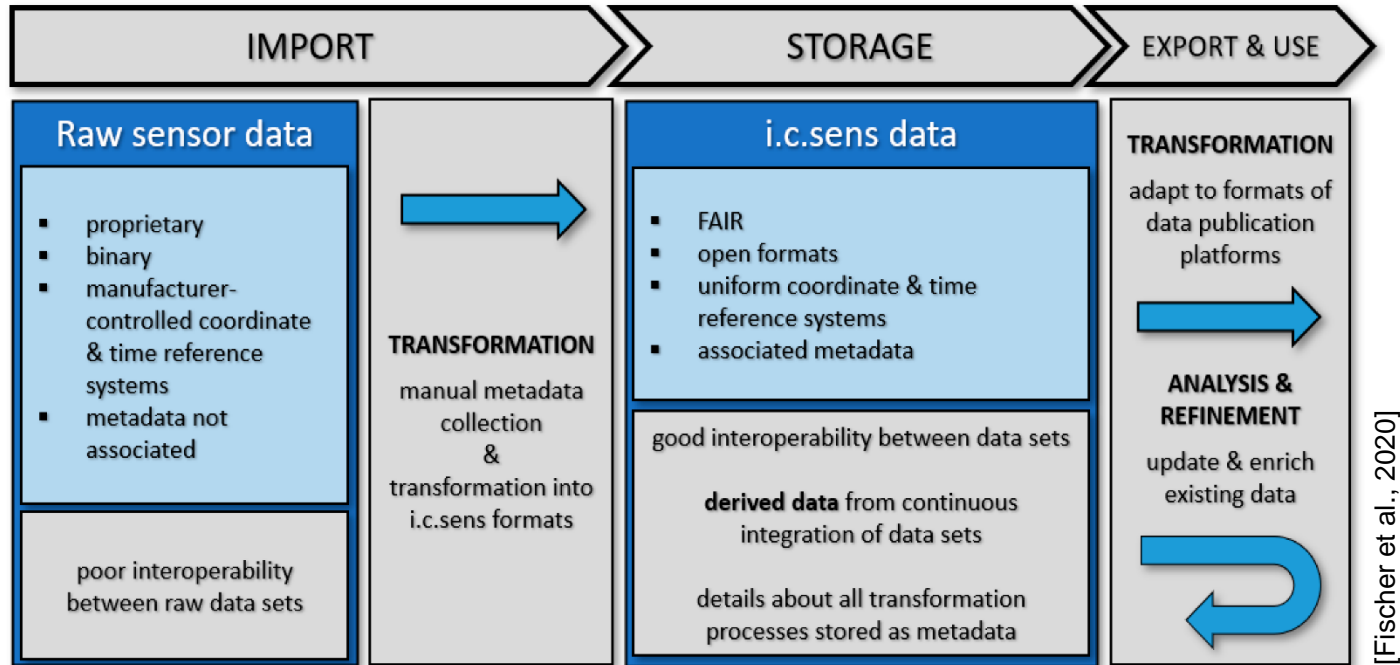


Lizenzen sind auch im Zusammenhang mit Forschungsdaten wichtig zu beachten und oftmals verpflichtend, wenn diese publiziert werden

Z.B. Creative Commons Licenses | creativecommons.org

- CC0 Universell → Keine Beschränkungen
- CC BY Namensnennung
- CC BY-SA Namensnennung - Weitergabe unter gleichen Bedingungen
- CC BY-ND Namensnennung - Keine Bearbeitung
- CC BY-NC Namensnennung - Nicht kommerziell





Schön et al. (2018)



Vogel (2020)

EXPERIMENT_1

SENSOR_PLATFORM_1

PLATFORM_CALIBRATION_DATA: includes data from platform calibration, i.e., raw measurements and obtained transformations between sensors

ROS_BAGS: storage for raw version of data logged by the ROS computer (includes data from stereo camera and GNSS/IMU system) as ROS bags; useable to re-create the sensor data traffic during recording

MMS

FULL_PROJECT: proprietary MMS storage format can be used by proprietary software to export various types of MMS-related data; this is considered the raw data for the MMS; metadata XML file

EXPORTED_POINT_CLOUD: point cloud data in various formats, e.g., colored point cloud with absolute coordinates separated into uniform spatial grid cells to reduce file size, full (down-sampled) point cloud; metadata XML file

EXPORTED_TRAJECTORY: export from MMS-internal GNSS as ASCII text file; metadata XML file

STEREO_CAMERA

STEREO_CAMERA_1

CAMERA_CALIBRATION_DATA: data from camera calibration, i.e., raw images and obtained (intrinsic) camera parameters

IMAGE_DATA: includes pairs of left/right images and an ASCII table that maps timestamps to image IDs; metadata XML file

GNSS/IMU

GNSS/IMU_1


PROPRIETARY_FORMAT: original sensor-dependent format, in some cases, only useable using sensor-specific proprietary software; metadata XML file

EXPORTED_FORMAT: export to accessible, interoperable ASCII format after export from the proprietary format using proprietary software; metadata XML file



Mögliche Ordnerstruktur für die gemessenen Sensordaten eines MSS

Fischer et al. (2020)


Leibniz
Universität
Hannover


[Datasets](#)
[Organizations](#)
[Groups](#)
[About](#)
[Legal](#)

[Organizations](#) / [i.c.sens](#) / [i.c.sens ...](#)

i.c.sens Visual-Inertial-LiDAR Dataset

Followers
1

Organization



i.c.sens
Graduiertenkolleg 2159
<https://www.icsens.uni-hannover.de/> [read more](#)

License
Creative Commons Attribution-NonCommercial 3.0

[Dataset](#) [Groups](#) [Activity Stream](#)

i.c.sens Visual-Inertial-LiDAR Dataset

The *i.c.sens Visual-Inertial-LiDAR Dataset* is a data set for the evaluation of dead reckoning or SLAM approaches in the context of mobile robotics. It consists of street-level monocular RGB camera images, a front-facing 180° point cloud, angular velocities, accelerations and an accurate ground truth trajectory. In total, we provide around 77 GB of data resulting from a 15 minutes drive, which is split into 8 rosbags of 2 minutes (10 GB) each. Besides, the intrinsic camera parameters and the extrinsic transformations between all sensor coordinate systems are given. Details on the data and its usage can be found in the provided [documentation file](#).

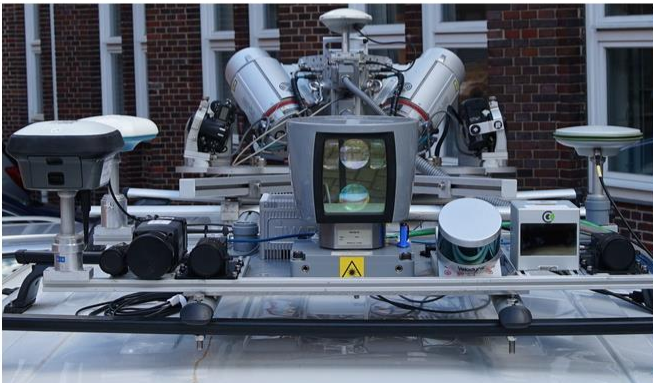





















Image credit: Sören Vogel

Data and Resources

- 
Vehicle
Vehicle used for our measurement campaign Image credit: Sören Vogel File size: 126.3 KByte
- 
Route
The route we took for our measurement campaign. File size: 634.8 KByte
- 
Sensor platform
The self-developed sensor platform containing several cameras, multiple... File size: 149.9 K
- 
Examples
Exemplary images and corresponding point clouds from our dataset. File size: 543.6 KByte
- 
Extrinsic transformation parameters
Extrinsic transformation parameters between the coordinate systems of the... File size: 1.0 K
- 
Intrinsic camera parameters
Intrinsic parameters of the camera including the camera matrix and five... File size: 285.0 Byte
- 
rviz configuration file
A configuration file for the 3D visualization tool for ROS: rviz. By... File size: 6.3 KByte
- 
Documentation 
Documentation explaining the data acquisition, the data format and the... File size: 20.8 MByte
- 
icsens-visual-inertial-lidar-dataset-1.bag 
The measurement data is split into 8 rosbags (2 minutes, 10 GB each). Refer... File size: 9.2
- 
icsens-visual-inertial-lidar-dataset-2.bag
The measurement data is split into 8 rosbags (2 minutes, 10 GB each). Refer... File size: 9.3
- 
icsens-visual-inertial-lidar-dataset-3.bag
The measurement data is split into 8 rosbags (2 minutes, 10 GB each). Refer... File size: 9.3
- 
icsens-visual-inertial-lidar-dataset-4.bag
The measurement data is split into 8 rosbags (2 minutes, 10 GB each). Refer... File size: 9.3
- 
icsens-visual-inertial-lidar-dataset-5.bag
The measurement data is split into 8 rosbags (2 minutes, 10 GB each). Refer... File size: 9.3
- 
icsens-visual-inertial-lidar-dataset-6.bag
The measurement data is split into 8 rosbags (2 minutes, 10 GB each). Refer... File size: 9.3
- 
icsens-visual-inertial-lidar-dataset-7.bag
The measurement data is split into 8 rosbags (2 minutes, 10 GB each). Refer... File size: 9.3
- 
icsens-visual-inertial-lidar-dataset-8.bag
The measurement data is split into 8 rosbags (2 minutes, 10 GB each). Refer... File size: 6.9

Cite this as

Raphael Voges (2020). Dataset: i.c.sens Visual-Inertial-LiDAR Dataset. <https://doi.org/10.25835/0026408> 

Retrieved: July 14, 2022, 10:28 AM (UTC+02:00)

- Soßna, V. und Ziedorn, F. (2020): Digitale Forschungsdaten managen - Grundlagen, Tipps und Tricks, Präsentation des Service-Teams Forschungsdaten, Technische Informationsbibliothek, Leibniz Universität Hannover.
- Research data managements. <https://www.uni-leipzig.de/en/research/research-service/research-data-management/>. Zugriff: 07.01.2021.
- Umgang mit Forschungsdaten. https://www.dfg.de/foerderung/antrag_gutachter_gremien/antragstellende/nachnutzung_forschungsdaten/index.html. Zugriff: 07.01.2021.
- The fair data principles. <https://www.force11.org/group/fairgroup/fairprinciples>. Zugriff: 07.01.2021.
- Richtlinie zum Umgang mit Forschungsdaten an der Leibniz Universität Hannover. <https://www.uni-hannover.de/de/universitaet/profil/ziele-strategien/umgang-forschungsdaten/>. Zugriff: 07.01.2021.
- Forschungsdatenmanagement. <https://www.fdm.uni-hannover.de/de/>. Zugriff: 07.01.2021.
- <https://rdmpromotion.rbind.io/promotion/> Zugriff: 07.01.2021.
- <https://www.tib.eu/en/publishing-archiving> (DOI-Service, Open Access). Zugriff: 07.01.2021.
- Fischer, C.; Sester, M.; Schön, S. Spatio-Temporal Research Data Infrastructure in the Context of Autonomous Driving. ISPRS Int. J. Geo-Inf. 2020, 9, 626.
- Raphael Voges (2020). Dataset: i.c.sens Visual-Inertial-LiDAR Dataset. <https://doi.org/10.25835/0026408>
- Max Coenen (2020). Dataset: i.c.sens Stereo Image Vehicle Dataset. <https://doi.org/10.25835/0078519>
- Uyen Nguyen (2020). Dataset: i.c.sens Multi-view pedestrian tracking dataset. <https://doi.org/10.25835/0082741>
- Schön, S.; Brenner, C.; Alkhatib, H.; Coenen, M.; Dbouk, H.; Garcia-Fernandez, N.; Fischer, C.; Heipke, C.; Lohmann, K.; Neumann, I.; Nguyen, U.; Paffenholz, J.-A.; Peters, T.; Rottensteiner, F.; Schachtschneider, J.; Sester, M.; Sun, L.; Vogel, S.; Voges, R. und Wagner, B. (2018): Integrity and Collaboration in Dynamic Sensor Networks. . In: Sensors. Schön, Steffen; Brenner, Claus; Alkhatib, Hamza; Coenen, Max; Dbouk, Hani; Garcia-Fernandez, Nicolas; Fischer, Colin; Heipke, Christian; Lohmann, Katja; Neumann, Ingo; Nguyen, Uyen; Paffenholz, Jens-André; Peters, Torben; Rottensteiner, Franz; Schachtschneider, Julia; Sester, Monika; Sun, Ligang; Vogel, Sören; Voges, Raphael; Wagner, Bernardo. 18 7, 21.
- Vogel, S. (2020): Kalman Filtering with State Constraints Applied to Multi-sensor Systems and Georeferencing. Ph.D. Thesis. München. DGK, Reihe C. 856.