Concept for data processing in the research project KOMET, in particular the service OPTIMAP, in compliance with data protection laws and Art. 5 GDPR

1. Who is conducting the research project or who is responsible for the specialist process / procedure?

   KOMET team members (researchers and developers) at the Chair of Geoinformatics, Department of Geosciences, TUD Dresden University of Technology.

2. For what explicit and legitimate purposes will the data be processed (Art. 5 para. 1 lit. b GDPR)?

   The objective of the service OPTIMAP is to improve findability of research outputs (papers, datasets, preprints) by collecting geospatial metadata (locations, time periods) alongside traditional publication metadata (title, author, publication venue, publication date, etc.) and making these metadata accessible via an interactive website. The increased findability and enhanced search improves access to research outputs for researchers, including across disciplines, the general public who can identify research covering their areas of interest, e.g., their home, and authors of research works, who benefit from exposure of their work.

3. What are the inclusion and exclusion criteria for test subjects or, in the case of business processes/procedures, which groups of people are affected by the data processing?

   All people who voluntarily create an account on OPTIMAP (optimap.geo.tu-dresden.de). These users consent to the data processing at the time of their first login.

   All authors of research outputs, e.g., journal articles or preprints, who publish work under their own name and personal information (affiliation, email, etc.). For this purpose, § 12 SächsDSDG applies: Data processing is permitted because it is necessary for the performance of scientific research, in particular the purpose of said research cannot be achieved with disproportionate effort and because the scientific interest in the performance of the research project outweighs the interest of the data subject in not having the data processed.

4. The following personal data shall be used in compliance with Art. 5 para. 1 lit. c GDPR:

   - Email
   - Publication metadata
     - Author name
     - Author affiliation
     - Author ORCID
   - Subscription area

As of: June 4, 2024

5. Legal basis for the processing of personal data

For the usage of emails as part of the user accounts for the platform, we explicitly ask for the user's consent. Users can delete their account themselves at any time.

For the collection (open data from third parties), processing, and provision of publication metadata § 12 SächsDSDG applies. The weighing of interests allow data storage and processing because it is necessary for scientific research, and any other means can only be achieved with disproportionate effort. The personal information that is part of the publication metadata was already published with the consent and in the interest of the involved parties (i.e., authors and co-authors), because they benefit from the publication of their work. The provision and distribution of said work in OPTIMAP further increases visibility of the authors' research output. Therefore, the data processing as part of the scientific research outweighs the interest of the data subject in not having the processing carried out.

The data collection as a third-party survey is possible based on the privilege regulation for research from Art. 14 para. 5 lit. a DGSVO. The information obligations shall not apply if and to the extent that the provision of such information proves impossible or would involve disproportionate effort, such as asking all authors named in publication metadata to consent explicitly. Therefore, we implement measures to safeguard the data subject's rights and freedoms and legitimate interests by making this data privacy concept and the privacy policy publicly available (https://optimap.geo.tu-dresden.de/privacy/). Data subjects continue to have a right to object, of course (see contact information in the privacy policy), which is checked on a case-by-case basis.

6. Implementation and data flow

All users can browse the OPTIMAP website and access the API freely without logging in. For specific features, such as notification about new articles for user-defined region, the user needs to create an account. The account consists solely of the email address stored in the application database. No password is stored. For each login, a login link with a login token is sent to user's email address. The token is valid for 10 minutes. By clicking on the link in the email, the user opens the website and is logged in. At the first login, the user must provide their consent to the privacy policy.

OPTIMAP is free and open source software based on Django, PostgreSQL and nginx. OPTIMAP is published under the GNU General Public License Version 3 (GNU GPL v3.0) at https://github.com/GeoinformationSystems/optimap.

7. Possible data protection considerations for individual items of the research project

Attribute	Data protection considerations	Justification	Measure
Email	Action needed	Highly individualized attribute	Access to server only for core team members, secure limited access to server, support for fully anonymous lo-

			gin via Email
Publication metadata: author name, affiliation, ORCID	Unproblematic	Already published metadata, individuals consent at the time of publication with respective publisher	
Subscription area	Unproblematic	Users create subscription areas themselves and can adjust them to not divulge a private location.	Alert the user at subscription creation about privacy implications.

8. <u>What technical and organizational measures are used to ensure privacy and data protection in accordance with Art. 5 para. 1 lit. e and f GDPR?</u>

   Data is collected in relation to individuals, which means it is theoretically possible to make inferences about individual persons based on the data collected. However, we guarantee that all information with privacy concerns are kept strictly limited to selected personal.

   The deployed application at https://optimap.geo.tu-dresden.de/ is only available via a secure HTTPS connection. The creation of tokens and secrets uses up-to-date cryptographic methods and the Django application is configured in a secure deployment mode. The PostgreSQL database runs in a Docker container in a seperated network with the containerised application and ginx webserver – only the application container can access the database. The secure ZIH Backup Service is used to regularly backup the server. Log files are purged regularly and in a production setting only contain information relevant for finding errors and maintenance of the service.

9. <u>Overall assessment of the risk of data processing for data subjects</u>

   The overall risk of data processing is low.
   The collected personal attributes are limited to the absolute minimum needed to offer the desired service and explicit consent is required.
   The collected third-party data (publication metadata) is already published with consent by the concerned individuals.