NFDI4Earth

Milestone MS1.3.9

Target group-specific curricula

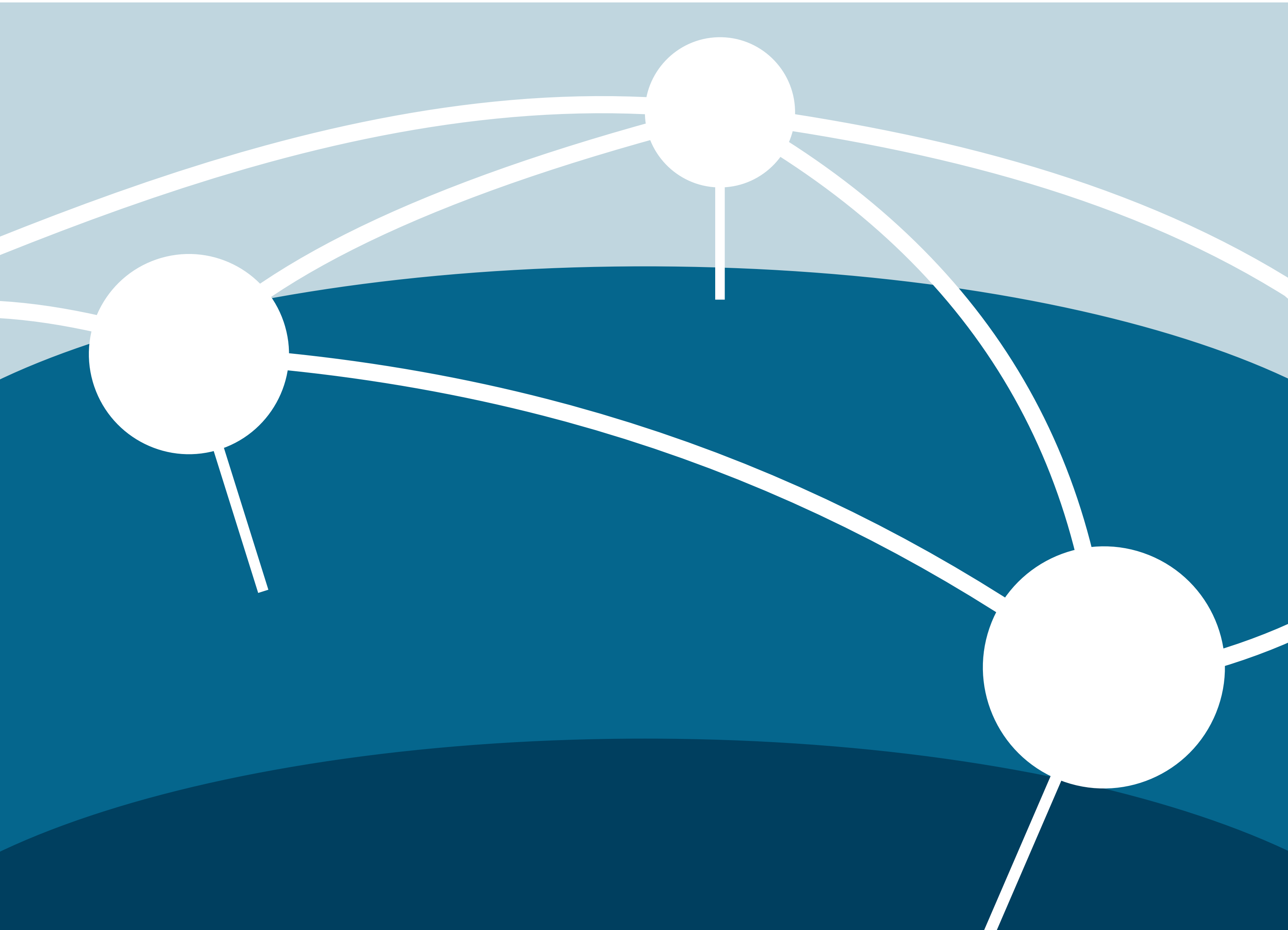
First revision

Farzaneh Sadeghi (farzaneh.sadeghi@hs-bochum.de), Carsten Keßler

2024-06 | Version: 1.0

DOI: 10.5281/zenodo.13768718

nfdi4earth.de



Citation

Farzaneh Sadeghi, Carsten Keßler. 2024. *Target group-specific curricula (NFDI4Earth Milestone MS1.3.9) (NFDI4Earth Deliverable MS1.3.9)*. NFDI4Earth Community on Zenodo. https://doi.org/10.5281/zenodo.13768718

License

Acknowledgement

This work has been funded by the German Research Foundation (DFG) through the project NFDI4Earth (DFG project no. 460036893, https://www.nfdi4earth.de/) within the German National Research Data Infrastructure (NFDI, https://www.nfdi.de/).

Executive summary

The NFDI4Earth EduTrain team is dedicated to supporting researchers in the Earth System Sciences (ESS) by providing comprehensive training resources. This report outlines the development and implementation of a tailored curriculum aimed at addressing the specific needs of the community. The curriculum development process was guided by a thorough needs assessment, identifying key challenges and requirements within the community. The resulting curriculum comprises 50 courses, primarily focused on spatial data analysis and management, with a strong emphasis on using open and reproducible methods. A modular approach was adopted to ensure flexibility and adaptability, allowing for the integration of new topics and technologies as they emerge. The courses were structured based on Bloom's taxonomy, categorizing them into cognitive levels to facilitate a clear and progressive learning pathway.Additionally, the curriculum incorporates suggestions of Open Educational Resources for each course to enrich the content. The NFDI4Earth curriculum represents a significant step towards building a robust educational framework for ESS researchers, fostering open and reproducible research practices.

Abbreviations

EDSDS - Earth System Data Science

ESS - Earth System Science

EduTrain - NFDI4Earth Education and Training team

KH - Knowledge Hub

NFDI - National Research Data Infrastructure

NFDI4Earth - National Research Data Infrastructure for Earth System Science

OER - Open Educational Resource

RDM - Research Data Management

Contents

1. Introduction

The NFDI4Earth project is committed to empowering Earth System Sciences (ESS) researchers by providing the tools and knowledge they need to manage and analyze their research data. As part of this mission, the project's Education and Training team (EduTrain) is developing a comprehensive suite of training resources designed to address the training needs of the ESS community. These resources aim to equip researchers with the skills and expertise necessary to use NFDI4Earth services and conduct open and reproducible research.

The primary target audience for the educational resources and services provided by EduTrain includes researchers at all career stages, from PhD students and postdoctoral researchers to senior scientists and professors. Additionally, the project aims to support ESS educators, training professionals, and university students pursuing ESS-related disciplines.

This report builds upon the findings from the needs assessment described in the preceding deliverable, "Mapping of existing educational resources and initial education and training needs within the Earth system science community" [1].

This report provides a detailed overview of the NFDI4Earth curriculum development process and the results designed for the designated target audience. Considering the rapidly evolving field of research data management, this curriculum emphasizes flexibility and adaptability. Its modular design facilitates the incorporation of new topics and technologies as they emerge. The curriculum will encompass a broad array of topics, ranging from the fundamentals of spatial data analysis and management to advanced techniques in data analysis using open and reproducible methods and tools.

In the following section, we briefly discuss the curriculum development models and tools, before we proceed to describe the NFDI4Earth curriculum development process. Subsequently, the results will be presented, followed by a discussion of future directions.

2. Background

Despite the frequent use of the term "curriculum" in educational research and policy, its precise definition continues to be a subject of ongoing debate [2]. In education, curriculum is comprehensively defined as the collective experiences that students encounter during the educational process. It includes the sequence of planned instruction and can encapsulate the deliberate interactions between students and instructional content, resources, and materials, as well as the processes involved in assessing educational objectives [3].

A curriculum should be both easily understandable and readily applicable. To remain relevant, it must continuously adapt to evolving educational values and expectations. Curriculum design involves organizing educational elements in a logical manner. There are two main types of curriculum design models: prescriptive models, which outline what curriculum designers should do, and descriptive models, which detail what curriculum designers actually do [4].

Prescriptive curriculum models offer guidelines for educators on curriculum design, emphasizing the establishment of clear, measurable objectives and the organization of educational experiences to meet these objectives. In contrast, descriptive curriculum models examine actual educational practices. These models consider the real-world contexts of teaching, adapting the curriculum to meet the specific needs and circumstances of the educational environment [4].

In this work, we have adopted a prescriptive curriculum design model for our online product. Since online platforms require clearly defined structures and objectives to effectively guide learners without the immediate feedback present in traditional classrooms, an adopted prescriptive approach is most suitable. This model allows us to establish explicit learning goals and organize content systematically, ensuring that the curriculum remains easily understandable and accessible to a diverse audience.

A crucial aspect of curriculum design is the articulation of learning objectives, which can be systematically classified using Bloom's Taxonomy [5]. Developed by educational psychologist Benjamin Bloom and his colleagues, this conceptual framework sorts the various cognitive levels of a subject. The revised version of the original taxonomy is currently used in numerous educational settings to categorize learning objectives. This revised taxonomy organizes learning objectives into hierarchical cognitive levels, spanning from fundamental skills such as remembering and understanding to more intricate abilities like evaluating and creating (see Fig. 1).

Structured curriculum models, along with tools such as Bloom's Taxonomy, are crucial for developing nurturing educational environments that support the development of a variety of cognitive skills and abilities.

In e-learning environments, structured approaches are becoming increasingly important. Users of online educational resources frequently lack immediate feedback or guidance, posing challenges to effectively navigating the content. This is particularly significant when users encounter numerous resources and need to identify the most relevant ones. Therefore, organizing digital learning resources into clear, manageable segments can significantly enhance their effectiveness. The implementation of structured curriculum models in the design of digital learning resources enhances focus and efficiency. This structure aids learners in achieving their educational objectives independently, even when they are separated from instructors and peers. Moreover, it makes it easier for educators to discover and reuse relevant materials [6].

Communicating curricular information through knowledge maps enables users to perceive

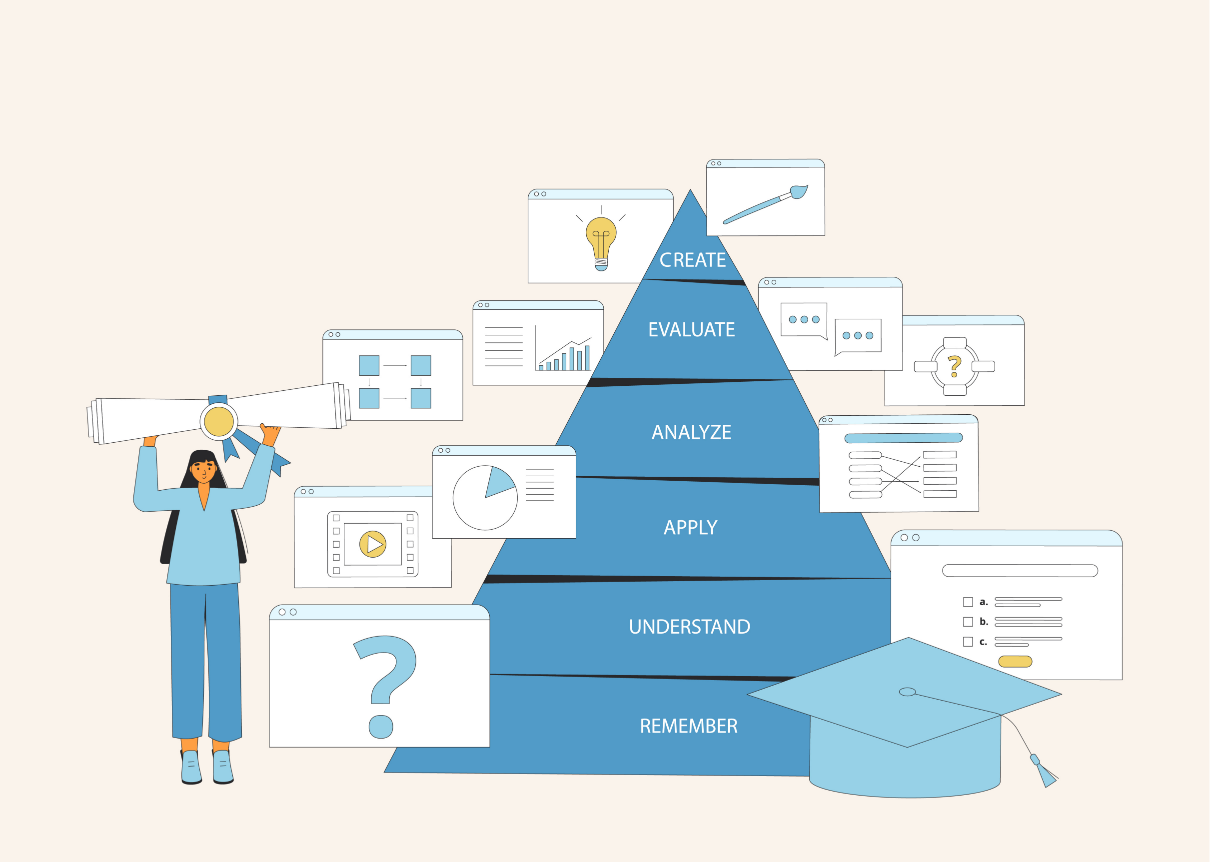


Figure 1: The Revised Bloom's Taxonomy [5]

the connections among various topics, thus preventing the issue of concentrating on isolated details without grasping the overall picture. Knowledge graphs and ontologies further enhance this process. Knowledge maps visually represent the interconnections between concepts. Knowledge graphs dynamically link these concepts, facilitating straightforward navigation. Ontologies define these relationships within a structured framework, ensuring precise knowledge organization. Collectively, these tools provide a comprehensive and navigable structure for curriculum information [7], [8].

A solid understanding of curriculum design, its diverse models, and associated tools will be instrumental as we develop the NFDI4Earth curriculum.

3. Building the NFDI4Earth Curricula

The NFDI4Earth curriculum is designed to encompass a wide range of topics essential for managing and analyzing spatio-temporal data using open and reproducible methods. To ensure that the curriculum remains adaptable and current, a modular approach has been adopted. This strategy directly responds to a significant challenge identified in the NFDI4Earth training needs assessment, highlighting the necessity for adaptable educational resources to effectively support ongoing research and data handling practices in the field [1].

Breaking down the subject areas into smaller, self-contained units, each focusing on a specific topic or skill, enhances the flexibility and customization of training delivery. Learners can choose the modules most relevant to their needs, interests, and previous knowledge, and progress through the material at their own pace. Additionally, educators can combine modules in various ways to create tailored training programs for specific audiences.

3.1. Data Collection

In developing the NFDI4Earth curriculum, we have extensively drawn from established curricula. Our foundational content is based on EarthLab's Introductory and Intermediate Earth Data Science textbooks [9], [10], which span a spectrum from fundamental to advanced methods of Earth data analysis and management using open and reproducible tools. This was complemented by the curriculum of the PhD course “Python for Geospatial Analysis” from Aalborg University [11], focusing on mapping, exploring, processing, and analyzing spatial information using entirely reproducible tools at an advanced level.

Additionally, the Module Handbooks from the Bachelor's and Master's programs in Geodesy and Geoinformation [12], [13], and the Master's program in Cartography [14] at the Technical University of Munich have been instrumental in expanding our data pool. These programs, which are highly relevant to the NFDI4Earth EduTrain initiative, share common areas of interest and have distributed some of their course material through an existing EducationalPilot.

We also drew inspiration from the EO4GEO Body of Knowledge to shape the learning objectives and establish connections between courses. EO4GEO delineates the interconnected concepts of Geographic Information and Earth Observation, which are essential to include in educational and training programs [15].

3.2. Data Preprocessing

We utilized Python to extract and compile metadata from EarthLab's textbook Markdown files (each textbook is composed of multiple lessons, each of which is available as a Markdown file) into a structured CSV format. This process involved parsing metadata, which included fine-grained learning objectives and descriptions for each topic, and storing the data in a CSV file. In total, 377 records were included in the final CSV file.

The lessons from the "Python for Geospatial Analysis" course have been processed and are now available online via the EduTrain portal. These lessons are not necessarily prerequisites for each other; rather, each lesson offers an overview of a broad array of tools essential for conducting reproducible spatial data analysis using Python. Utilizing the EO4GEO framework, we defined learning objectives and prerequisite knowledge for each lesson. In total, 11 advanced topics were identified.

The module handbooks from the Technical University of Munich are available as PDF files. A custom PDF scraping algorithm was developed to facilitate data extraction and convert selected PDF content into structured CSV format. This algorithm specifically extracts the title, description, learning objectives, and recommended prerequisites for each course. In total, 182 records were included in the final CSV file.

3.3. Sketching the Curriculum

For the initial version of the NFDI4Earth curriculum, prerequisites were chosen as the main feature to define the relationships between course entities.

Initially, the following courses from our portal were identified within the data pool. These courses, developed by our EduPilots and collaborators, provide advanced-level educational content:

- Fundamentals of Spatial Data Mining and Machine Learning
- Image Processing and Analysis
- Lessons in Python for Spatial Data Analysis

We then located the prerequisites defined by the programs to which each of these courses belongs and performed a similar analysis for those prerequisites, continuing until we reached the most introductory topics, such as “Introduction to the Science of Geographic Information Systems” and “Setting Up a Python Earth Data Analysis Environment.” This process resulted in the first draft of the curriculum. Subsequently, the relationships among the courses were refined.

After extracting this initial set of interconnected courses from the course pool, the courses of the curriculum and their attributes were incrementally refined and shaped. The following sections will provide details on the refining process.

3.3.1. Learning Objectives

This refinement process began with an analysis of the learning objectives for each course. We revised the objectives for courses where the original objectives were unclear or included non-reproducible practices. This was achieved with the assistance of the EO4GEO knowledge map and a thorough review of literature and textbooks covering spatial data analysis using open tools, primarily the Earth Data Science textbooks by EarthLab [9], [10].

3.3.2. Course Difficulty

A major challenge in defining prerequisite relationships and constructing a system of interconnected courses rather than a linear program was the courses’ varying difficulty levels. For creating a modular, interconnected curriculum—essentially a collection of numerous learning pathways—simply categorizing a course as master’s or bachelor’s level does not sufficiently describe its complexity. This classification system is logical for academic institutions where an admission process often includes suggesting or mandating prerequisites. However, when dealing with open educational resources in Earth System Sciences (ESS) that cover a broad range of topics, making assumptions about the background knowledge of users is impractical.

To address this challenge, each course was assigned a cognitive level by mapping the revised learning objectives to the cognitive domains defined in Bloom’s taxonomy. The association of courses with cognitive levels was based on the most complex topic each course covers (Fig. 1).

This mapping was then used to assign a difficulty level to each course, as detailed below in Tab. 1.

Table 1: Guidelines for standardizing the difficulty levels of NFDI4Earth courses

Cognitive Level	Cognitive Domain	Course Level
Level 1	Remember	beginner
Level 2	Understand	beginner
Level 3	Apply	intermediate
Level 4	Analyze	intermediate
Level 5	Evaluate	advanced
Level 6	Create	advanced

Tab. 1 illustrates these classifications, ensuring that each course's complexity is accurately represented, facilitating learners in selecting the most appropriate modules for their skill level and educational needs.

To prevent the creation of overly lengthy courses covering a broad range of topics—which could overwhelm learners, courses that spanned multiple difficulty levels were divided into smaller segments. For each segment, we defined specific objectives and assigned a difficulty level as previously described. This division allowed us to establish more precise relationships between courses, transitioning from a linear system, where a large course was a prerequisite for another, to a modular and networked system. In this network, we mapped specific parts of formerly large courses to corresponding parts of others, transitioning from a strict prerequisite relationship to a more nuanced prerequisite and progressive relationship. This progressive model moves learners from beginner to advanced levels within a topic.

Additionally, we ensured that each ESS topic began at a beginner level and progressed to more advanced stages. For topics lacking any of the three steps, we defined new courses, drawing inspiration from our data pool, open university textbooks, and other resources. This iterative refinement process was repeated to ensure all courses were integrated seamlessly into the curriculum.

3.3.3. Course Title

We aimed to maintain a coherent naming pattern based on Bloom's taxonomy. While course learning objectives and difficulty levels are crucial for clear communication with users, the course titles should also provide descriptive and useful information whenever possible, such as indicating the level and the tool used.

Table 2: Approaches for naming NFDI4Earth Courses

Cognitive Level	Cognitive Domain	Naming Pattern	Course Example
Level 1	Remember	Introduction to	Introduction to Python packages for Earth data science
Level 2	Understand	Introduction to/ Fundamentals of	Fundamentals of geopandas
Level 3	Apply	Fundamentals of/ Work with	Work with vector data in Python
Level 4	Analyze	Analyze	Analyze vector data in Python
Level 5	Evaluate	Evaluate	Evaluate spatial data in Python

Tab. 2 does not cover the 6th cognitive level, "Create," which involves projects and independent research and is beyond the scope of our current service. It is important to note that this naming pattern was primarily used for courses focusing on ESS-specific tools and data types where the entire skill acquisition path was mapped. For more generic courses or complex topics, such as "Setting up a Python Earth Data Analysis Environment" or "Mathematics for Earth Data Science," this convention was not applied.

3.3.4. Learning Resources

Whenever possible, we linked each course to Open Educational Resources (OERs) that cover similar topics. Currently, most of these suggestions are sourced from the "Python for Spatial Data Analysis" lessons [11], chapters from EarthLab's textbooks [9], [10], our EduPilots, and our OER collection [1]. Course definitions and their interconnections form the skeleton of our curriculum, while OERs provide the necessary substance. Mapping existing OERs into the curriculum simplifies the process of reusing existing material and developing our own resources when necessary. This integration helps populate our educational portal, where users can interact with the content and evaluate the effectiveness of the suggested curriculum.

3.4. Constructing the Curriculum

After iterating the aforementioned processes several times and evaluating the identified relationships connecting the courses, the desired knowledge graph for constructing the NFDI4Earth curriculum was formulated.

We utilized Python to automate the transformation of data from CSV format into an RDF graph. Predefined namespaces from Schema.org and LRMI were employed to categorize the data consistently. Prerequisite and related resource links were dynamically generated and associated with each course, as detailed in Appendix 1. The final graph, serialized in Turtle syntax, represents each course as a learning resource annotated with titles, learning objectives, and difficulty level. Each course is also dynamically linked with prerequisite learning resources and related OER. The graph comprises a total of 380 triples.

4. The NFDI4Earth Curricula

The first version of the NFDI4Earth curriculum includes 50 courses, primarily focused on spatial data analysis and management, with a strong emphasis on Python-based learning. These courses were compiled by collecting and combining relevant educational resources from various reliable sources, creating a comprehensive dataset. We revised and standardized the learning objectives of these courses to align with our goals to provide resources on conducting open and reproducible science, ensuring consistency and relevance. The curriculum is intricately structured, as depicted in the following visualization, which maps out the courses and their interconnections (Fig. 2).

Fig. 2 Fig. 2 illustrates a comprehensive network of courses, showing a dense interlinking of topics that cover foundational to advanced aspects of spatial data handling and analysis. Each node represents a course, and the connecting lines represent prerequisite relationships, indicating a well-thought-out progression designed to build expertise systematically.

Fig. 3 provides a focused view of a subset of the curriculum, highlighting specific pathways such as the progression from introductory Python courses to more specialized applications in spatial data mining and machine learning. This map demonstrates the curriculum's modular yet integrated approach, allowing for both broad and deep exploration of the subjects.

Detailed information and the initial version of the curriculum are included in the attached file `ETKG_20240916.ttl`.

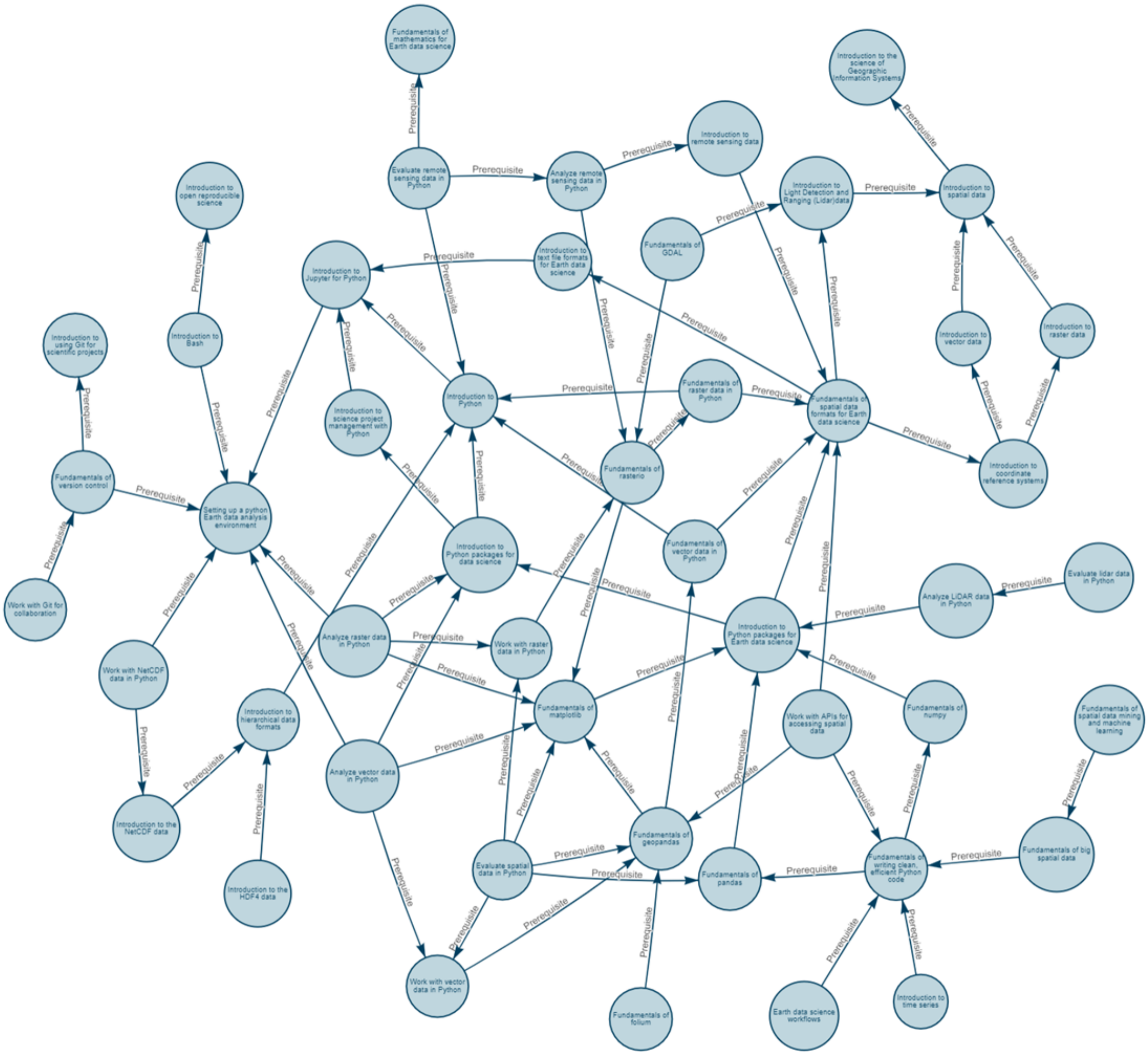


Figure 2: Visualization of NFDI4Earth curriculum

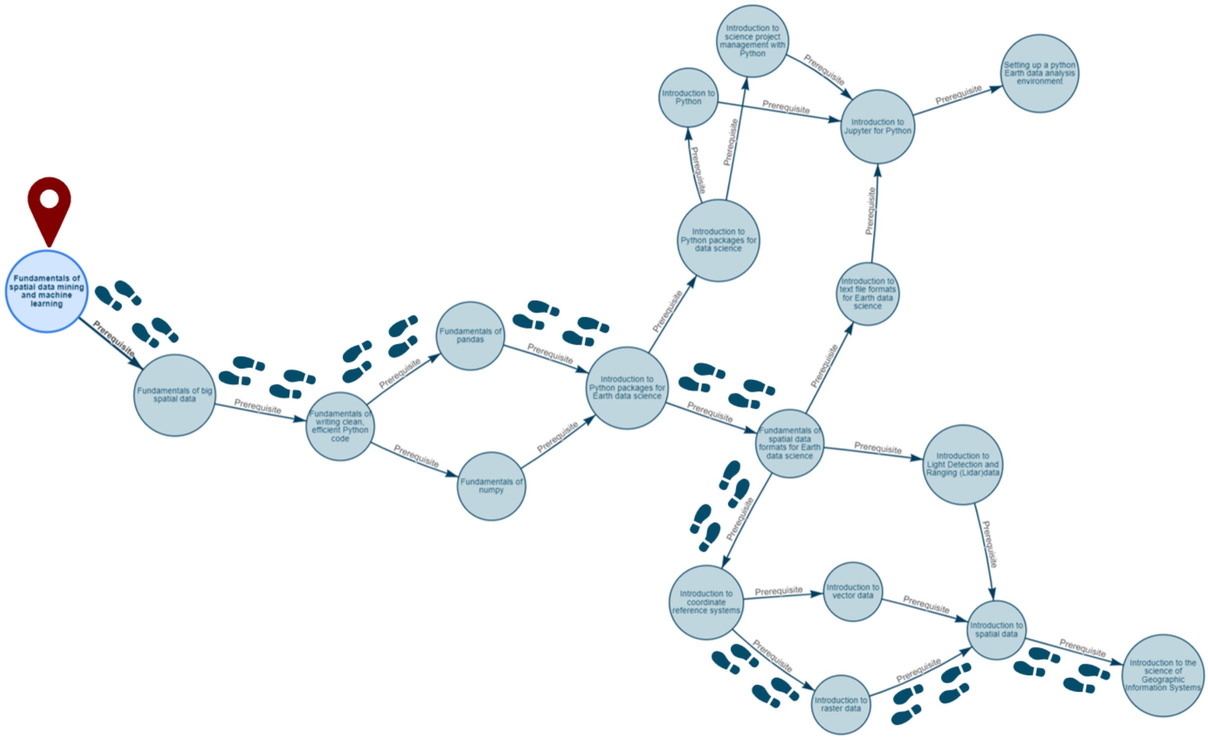


Figure 3: Learning pathways

5. Conclusion and Future Direction

The NFDI4Earth curriculum aims to deliver a thorough educational plan in spatial data analysis using open and reproducible methods. It is designed to help learners acquire the skills needed to manage spatial data in research and professional contexts. The curriculum dataset outlines prerequisites for each course, allowing learners to understand the required prior knowledge and customize their learning paths based on their skill levels.

The educational material for the courses outlined in the curriculum will be gradually published. Some of the course material is already available on the EduPortal, and efforts are ongoing to develop and publish quality educational content covering the concepts of this curriculum. These materials will be released in the upcoming phases of the project. All published course information will be integrated into KnowledgeHub to improve accessibility and broaden reach.

The future plans for the project include expanding the range of topics covered and exploring cross-cutting themes to enrich the curriculum. There is also a focus on incorporating R for data analysis. An important advancement will be the implementation of automated learning pathway generation, which will customize educational paths based on individual learning needs and objectives, enhancing the educational experience for all users.

6. References

[1] F. Sadeghi *et al.*, "Mapping of existing educational resources and initial education and training needs within the Earth system science community (NFDI4Earth Deliverable D1.3.1)," May 2023, doi: 10.5281/ZENODO.7940195.

[2] M. Young, "What is a curriculum and what can it do?," *The Curriculum Journal*, vol. 25, no. 1, pp. 7–13, Mar. 2014, doi: 10.1080/09585176.2014.902526.

[3] A. V. Kelly, *The curriculum: theory and practice*, 6th ed. Thousand Oaks, CA: SAGE Publications, 2009.

[4] D. Prideaux, "ABC of learning and teaching in medicine: Curriculum design," *BMJ*, vol. 326, no. 7383, pp. 268–270, Feb. 2003, doi: 10.1136/bmj.326.7383.268.

[5] L. W. Anderson and D. R. Krathwohl, Eds., *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*, Complete ed. New York: Longman, 2001.

[6] H. Zhu *et al.*, "A multi-constraint learning path recommendation algorithm based on knowledge map," *Knowledge-Based Systems*, vol. 143, pp. 102–114, Mar. 2018, doi: 10.1016/j.knosys.2017.12.011.

[7] P. Pei, D. Xuejing, and Z. De"ing, "Construction of Curriculum Knowledge Map based on Ontology," in *Proceeding" of the 2020 2nd International Conference on Robotics, Intelligent Control and Artificial Intelligence*, in RICAI '20. New York, NY, USA: Association for Computing Machinery, Dec. 2020, "p. 259–265. doi: 10.1145/3438872.3439091.

[8] X. Yu, M. Stahr, H. Chen, and R. Yan, "Design and Implementation of Curriculum System Based on Know'20ge Graph," in *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, Jan. 2021, pp. 767–770. doi: 10.1109/ICCECE51280.2021.9342370.

[9] "Introduction to Earth Data Science Earth"Lab CU Boulder," Earth Data Science - Earth Lab. Accessed: Jun. 23, 2024. [Online]. Available: https://www.earthdatascience.org/courses/intro-to-earth-data-science/

[10] "Use Data for Earth and Environme"tal Science in Open Source Python Earth Lab CU Boulder," Earth Data Science - Earth Lab. Accessed: Jun. 23, 2024. [Online]. Available: https://www.earthdatascience.org/courses/use-data-open-source-python/

[11] C. Keßler and J. Jokar Arsanjani,""Python for Geospatial Analysis (2023)." Jun. 15, 2024. Accessed: Jun. 23, 2024. [Online]. Available: https://github.com/crstn/Geospatial-Analysis-2023

[12] Technical University of Munich, *Module catalog b.s". Geodesy and geoinformation*. 2021. ["nline]. Available: https://collab.dvb.bayern/pages/viewpage.action?pageId=73390039&preview=/73390039/73392000/Module-Catalog_BScGeodesyGeoinformation_EN_01122021.pdf

[13] Technical University of Munich, *Module catalog m.sc. Geodesy and geoinformation*. 2022. [Online]. Available: https://collab.dvb.bayern/pages/viewpage.action?pageId=73390039&preview=/73390039/73392000/Module-Catalog_BScGeodesyGeoinformation_EN_01122021.pdf

[14] Technical University of Munich, *Module handbook m.sc. Cartography*. 2022. [Online]. Available: https://collab.dvb.bayern/pages/viewpage.action?pageId=73390039&preview=/73390039/73392000/Module-Catalog_MScGeodesyGeoinformation_EN_19102022.pdf

[15] EO4GEO Consortium, "EO4GEO - towards an innovative strategy for skills development and capacity building in the space geo-information sector supporting copernicus user uptake." [Online]. Available: http://www.eo4geo.eu/

A. Appendix 1

- `etkg`: A temporary custom namespace used for specific identifiers for each course, will be updated for integration into the KnowledgeHub.
- `schema.LearningResource`: This class is used to define each course as a learning resource.
- `schema.educationalLevel`: Used to describe the cognitive domain and difficulty level of the courses
- `schema.competencyRequired`: Specifies the recommended prerequisites for undertaking a course, linked to other course URIs.
- `DCTERMS.title`: Represents the name of the course.
- `DCTERMS.source`: Used to associate courses with external or supplementary learning resources.
- `lrmi.learningObjective`: Captures the course's learning objectives.