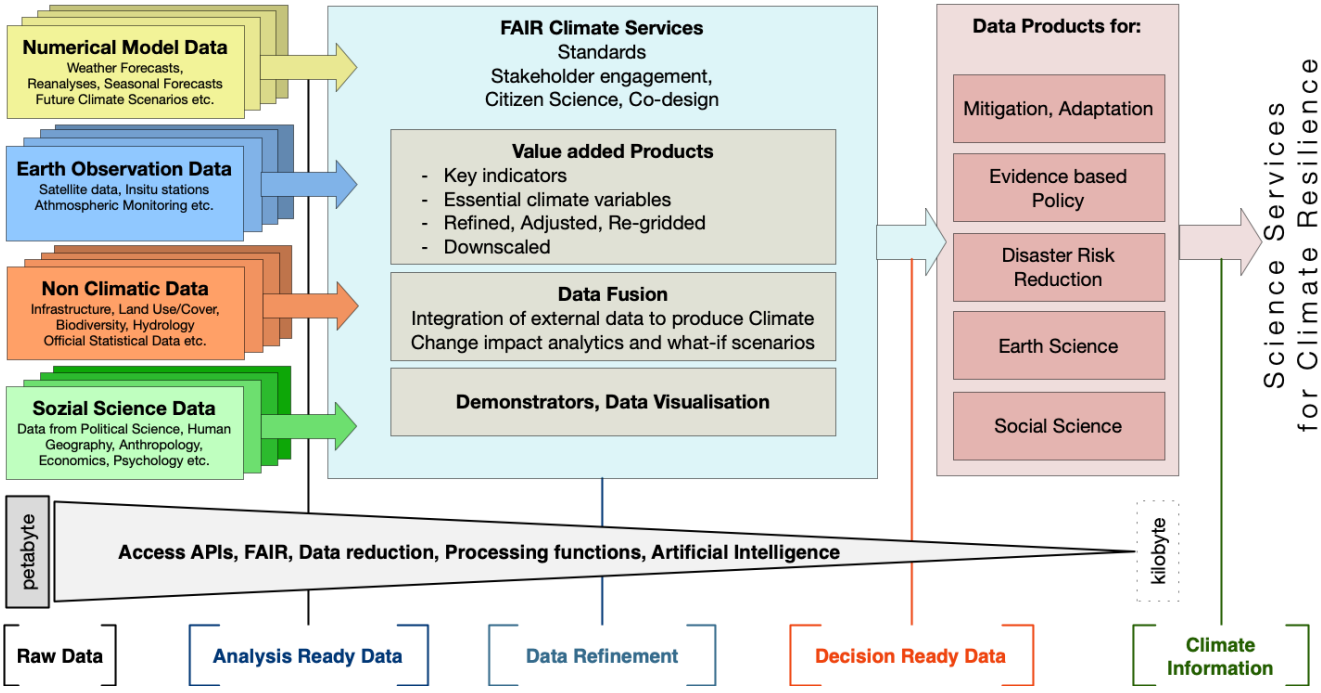
Enhancing Interoperability for Climate Change Services

Version 2.1 /18 Oct 2021

Chapter 1. Join us to Accelerate Climate Change Data Interoperability and Integration with Numerical Model Data, Earth Observation Data, Social Science Data and non Climatic Data.

The Open Geospatial Consortium is offering sponsorship opportunities to support the OGC Climate Change Service Pilots. The initiative will start in 2022 and run for 5 years with multiple focus areas and cycles where specific technical challenges will be worked on. The objective of the Climate Change Services Pilots is to accelerate our collective readiness for accessing, fusing, and analyzing data from the climate change modeling community with earth observation and sozial science data in order to contribute to the global push for achieving climate resilience. Our goal is to develop a reliable foundation for science services for climate change actions. For this purpose, OGC members engaged in this pilot will develop a series of demonstrators that show the integration and combined exploitation of numerical data, EO data, and data from sozial science. The following figure provides an overview of the envisioned pilot processes.



As illustrated, big, raw data from multiple sources requires further processing in order to be ready for analysis and climate change impact assessments. Applying data enhancement steps, such as bias adjustments, re-gridding, or calculation of climate indicators and essential variables, leads to “Decision Ready Climate Data.” The spatial data infrastructures required for this integration should be designed with interoperable building blocks following FAIR data principles. Heterogeneous data from multiple sources can be enhanced, adjusted, refined, or quality controlled to provide Science Services data products for Climate Resilience. The OGC Climate Change Services Pilots will also illustrate the graphical exploration of the Decision Ready Climate Data. It will demonstrate how to design FAIR climate services information systems. The OGC Pilot demonstrators will illustrate the necessary tools and the visualisations to address climate actions moving towards climate resilience.

By sponsoring the OGC Climate Change Service Pilot, organizations can collaborate on building the foundation for climate and non-climate data integration processes based on standardized APIs and resource models. The pilot demonstrators will unveil interoperability and integration issues, help to define the essential data processing functions, and establish a practical, repeatable, documented, community-driven and -vetted step away from existing, disconnected platforms and towards integration-ready services. To get in contact, please use the OGC Innovation Program contact form. **Response period for the first focus area ends December 31st, 2021**.

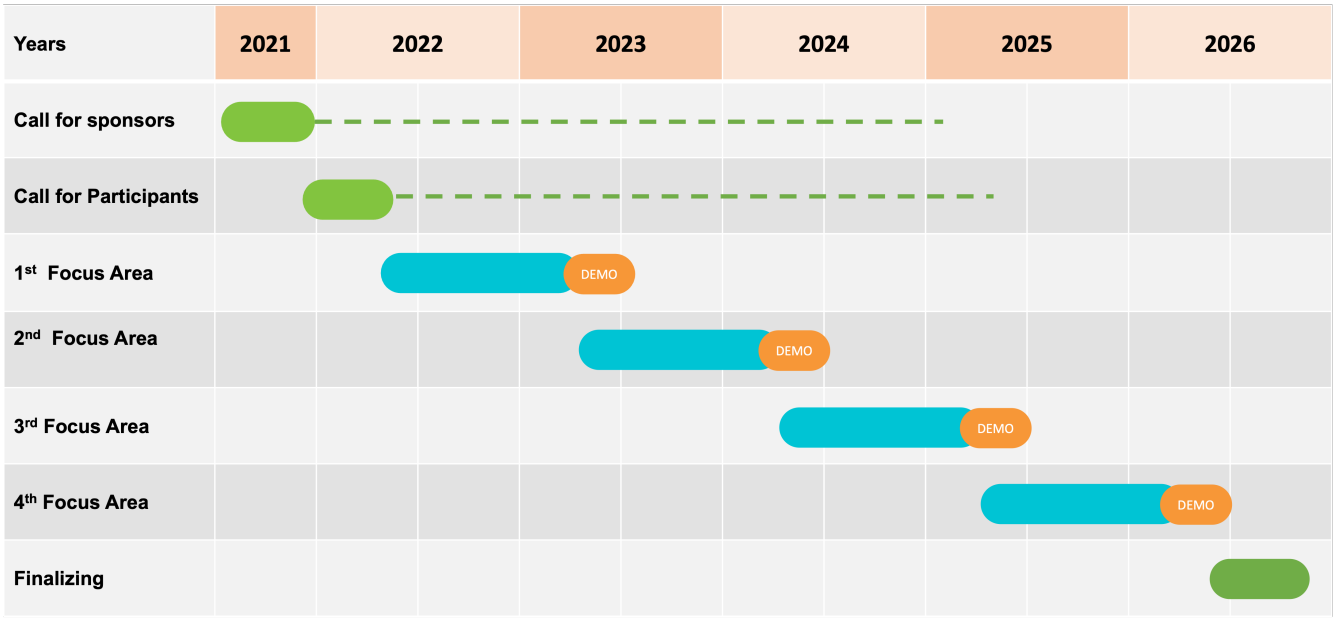
Chapter 2. Background

OGC technologies already play an essential role in the context of climate data discovery, exchange, and analysis. As an example, the ESA Climate Change Initiative (CCI) implements steps necessary to generate a consistent set of Essential Climate Variables (ECVs). With OGC Web Map Service (WMS) and Web Coverage Services (WCS) for data access and visualisation, instances of the OGC Catalogue Service for the Web (CSW) to express metadata for all CCI data stored in the CCI data archive, the CCI builds on OGC technologies to satisfy Global Climate Observing System's (GCOS) monitoring principles on data systems. CCI's products are then further developed into operational systems by non-ESA entities including the Climate Data Store (CDS) of the Copernicus Climate Change Service (C3S), EUMETSAT and its Satellite Applications Facilities, and other climate service initiatives with their data integration and analysis systems (e.g., the Data Integration and Analysis Systems, DIAS, and the emerging Earth Environmental Information Platforms).

Members of the climate change and impact community have prompted OGC to support the community with its standards, services, and best practices. They identified the opportunity for doing this most effectively is to target the interface between climate and other domains rather than seeking to model the full complexity within the climate domain itself. OGC standards provide a "lift up" towards the development of APIs for services by reducing the effort needed to establish community consensus while accelerating their adoption and impact.

Chapter 3. Timeline

The following figure illustrates the tentative timeline. The OGC Team will work with sponsoring organizations to adjust the timeline based on funding and timing constraints.



The Call for Sponsors for the 5 year Pilot remains open until December 31st, 2021. It is foreseen that the activity will run over 4-5 years with multiple thematic focus areas. Additional sponsors are therefore able to join the Climate Change Services Pilot after the first focus area cycle has started. After the Call for Sponsors period, OGC will, together with the sponsoring organizations, develop the Call for Participation for release before the end of 2021. Similar to the Call For Sponsors, the Call For Participation will allow new participants to join the activity at suitable time slots across the focus areas. The Call for Participation will describe the exact requirements to be addressed in the focus areas and will be launched publicly. OGC member organizations are eligible for responding to the call and will be invited to submit their proposals by mid-February, 2022. OGC and the sponsoring organizations select the organizations that best fit, based on their experience, products, and ability to meet the objectives. We expect to organize a kick-off meeting in May, 2022. The pilot will tentatively end September, 2022, with the results being presented at the OGC Member Meeting that quarter. As said, this is a tentative timeline and can be adapted to sponsors' needs.

Chapter 4. Technical Challenges

Realizing the delivery of Decision Ready Data on demand to achieve Climate Resilience involves a number of technical challenges that have already been identified by the community. A subset will be selected and embedded in use-cases that will be defined jointly by Pilot Sponsors and the OGC team. The goal is to ensure a clear value-enhancement pipeline as illustrated in Figure 1, above. This includes, among other elements, a baseline of standardised operators for data reduction and analytics. These need to fit into an overall workflow that provides translation services between upstream model data and downstream output - basically from raw data, to analysis-ready data, to decision-ready data. The following technical challenges have been identified and will be treated in the focus areas cycles of the Pilot accordingly:

- **Big Data Challenge:** Multiple obstacles still exist, creating big barriers for seamless information delivery starting from **Data Discovery**. Here the emergence of new data platforms, new processing functionalities, and thus new products, data discovery remains a challenge. In addition to existing solutions based on established metadata profiles and catalog services, new technologies such as OGC's Spatio-Temporal Asset Catalog (STAC) and open Web APIs such as OGC API Records will be explored. Furthermore, aspects of **Data Access** need to be solved where the new OGC API suite of Web APIs for data access, subsetting, and processing are currently utilized very successfully in several domains. Several code sprints have shown that server-side solutions can be realized within days and clients can interact very quickly with these server endpoints, thus development time is radically reduced. A promising specialized candidate for climate data and non-climate data integration has been recently published in the form of the OGC API - Environmental Data Retrieval (EDR). But which additional APIs are needed for climate data? Is the current set of OGC APIs sufficiently qualified to support the data enhancement pipeline illustrated in Figure 1? If not, what modifications and extensions need to be made available? How do OGC APIs cooperate with existing technologies such as THREDDS and OPEnDAP? For challenges of data spaces, **Data Cubes** have recently been explored in the OGC data cube workshop. Ad hoc creation and embedded processing functions have been identified as essential ingredients for efficient data exploration and exchange. Is it possible to transfer these concepts to all stages of the processing pipeline? How to scale both ways from local, ad hoc cubes to pan-continental cubes and vice versa. How to extend cubes as part of data fusion and data integration processes?
- **Cross-Discipline Data Integration:** Different disciplines such as Earth Observation, various social science, or climate modeling use different conceptual models in their data collection, production, and analytical processes. How can we map between these different models? What patterns have been used to transform conceptual models to logical models, and eventually physical models? The production of modern Decision-ready information needs the integration of several data sets, including census and demographics, further social science data, transportation infrastructure, hydrography, land use, topography and other data sets. This pilot cycle uses 'location' as the common denominator between these diverse data sets and works with several data providers and scientific disciplines. In terms of **Data Exchange Formats** the challenge is to know what data formats need to be supported at the various interfaces of the processing pipeline? What is the minimum constellation of required formats to cover the majority of use cases? What role do container formats play? Challenging on technical level is also the **Data Provenance**. Many archives include data from several production cycles, such as IPCC AR 5 and AR 6 models. In this context, long term support needs to be realized and full

traceability from high level data products back to the original raw data. Especially in context of reliable data based policy, clear audit trails and accountability for the data to information evolution needs to be ensured.

- **Building Blocks for processing pipelines:** With a focus on **Machine Learning and Artificial Intelligence** which plays an increasing role in the context of data science and data integration. This focus area needs to evaluate the applicability of machine learning models in the context of the value-enhancing processing pipeline. What information needs to be provided to describe machine learning models and corresponding training data sufficiently to ensure proper usage at various steps of the pipeline? Upcoming options to deploy ML/AI within processing APIs to enhance climate services are rising challenges e.g. on how to initiate or ingest training models and the appropriate learning extensions for the production phase of ML/AI. Heterogeneity in data spaces can be bridged with **Linked Data and Data Semantics**. Proper and common use of shared semantics is essential to guarantee solid value-enhancement processes. At the same time, resolvable links to procedures, sampling & data process protocols, and used applications will ensure transparency and traceability of decisions and actions based on data products. What level is currently supported? What infrastructure is required to support shared semantics? What governance mechanisms need to be put in place?
- **Visualization:** Science Service in general needs to deliver the information in usable form. Climate service in particular faces the challenge to visualize very complex statistical graphs which need to be displayed on demand within a short time at interfaces still operable under low bandwidth conditions.

Chapter 5. OGC Innovation Program

The OGC Innovation Program is an innovative, collaborative, and hands-on engineering and rapid prototyping program. In the Innovation Program, OGC members bring forward technology and technology integration challenges. These challenges are refined and mapped to a set of requirements, use cases, and implementation scenarios and eventually addressed in different types of initiatives. These initiatives bring OGC vendors and research institutions together with sponsoring organizations. Coordinated and managed by the OGC Innovation Team, each initiative has the goal to stepwise increase Technology Readiness Levels (TRL) for domain-driven solutions, including software architecture, interface design, information and data models, as well as related standards, specifications, and best practices. Run globally, the Innovation Program further validates and tests geospatial technology and innovations based on OGC standards and influences future OGC strategic work items.

This OGC Climate Change Pilot initiative will be conducted according to the OGC Innovation Program policies and procedures. An OGC Pilot unites technology users with technology solution providers in a fast-paced, collaborative prototyping environment to test and evaluate implementation of existing or emerging standards and best practices. Pilots serve as an incubator to prioritize, encourage, and accelerate the pace at which standards-based capabilities are deployed in order to improve interoperability within a specific domain or user community.

Chapter 6. Get in Contact

If you are interested in joining this exciting opportunity to contribute to our collective efforts to address climate resilience by enhancing the integration potential among our current climate services information systems, please contact the OGC Innovation Program via the OGC Innovation Program contact form. **Response period for the first focus area ends December 31st, 2021**.