24S651031: Copernicus Hubs and Institutions

Excursion Report

This report is a summary and reflection on the excursion that took place on the **June 10, 2025** in Vienna as part of the Copernicus Hubs and Institutions course**.** The visit was both an introduction and immersion to the operations and projects of the key institutions in the Earth Observation domain: Earth Observation Data Centre (EODC) and Environment Agency Austria.

Location: TU Wien and Umweltbundesamt offices

Agenda:

1. Morning session:
    a. Introduction to Earth Observation Data Centre by **Christian Briese**, CEO EODC
    b. Guided tour of the Vienna Scientific Clusters (VSC4, VSC5), MUSICA and Tape room.
2. Afternoon session:
    a. Introduction to the Environment Agency Austria by **Dipl.-Ing. Gebhard Banko** and **Roland Grillmayer** of the Remote Sensing and Spatial Analysis team at UBV.

Objectives:

Course objectives:

1. Locate key institutions in the EO/Copernicus domain within Europe.
2. Remember key actors and their contribution to the success of Copernicus.
3. Understand EU inspire and Copernicus data space ecosystem.
4. Relate various Copernicus components (upstream, downstream etc) with relevant institutional settings.

Personal objectives:

1. Understand the contribution of EODC and UBV in the EU Copernicus Data Space Ecosystem.
2. How EODC deals with Big Data and cloud computing in EO domain.
3. Understand how environment agency Austria ensures their data is open and accessible to the public.

Earth Observation Data Centre (EODC)

Introduction

The domain of Earth Observation (EO) has seen a continuous increase in data in terms of volume, variety, and complexity (European Commission. Joint Research Centre., 2023). The constant stream of EO data producing about 16TB (~ 4PB per year) of data filling archives every day and this increasing even further due to the availability of more satellites (Backeberg et al., 2023; Sudmanns et al., 2023). This excursion aimed to explore how Austria is addressing this challenge by visiting the Earth Observation Data Centre (EODC) and understanding its role in efficiently managing, processing, and storing the large amount of EO data.

Overview of EODC

EODC, founded in 2014, plays and important role in the distribution and processing the vast amount of EO data available today in Austria(Wagner et al., 2014). Its mission is to set-up and operate joint IT infrastructure with the public and private sector and to be a bridge between science and application. It provides different computing and storage systems.

During the visit, at TU Wien, we had an interactive session led by **Christian Briese,** the CEO EODC, who introduced the mission of EODC, its operations, projects and future advancements. The EODC IT infrastructure includes the following components: EODC Cloud, EODC Storage and EODC High Performance Computer (HPC); EODC processor and EODC backup which are exclusively utilized by EODC staff; EO Data Ingestion. EODC also provides satellite imagery Copernicus' Sentinel 1, 2 and 3 and Analysis Ready Data (ARD) from projects with Copernicus and ESA including Copernicus global land cover, Global flood monitoring service, interTwin, GREAT, ESA climate change initiative (CCI) and more.

Immediately thereafter, in 2 separate groups, we had a guided tour of the **Vienna Scientific Computer (VSC).** The VSC, started as a project of three universities in 2009, is a collaboration of several universities and research organizations today that provide supercomputer resources and services(European Commission. Joint Research Centre., 2023). The VSC is considered the largest and most powerful computer in Austria. Today there is the VSC 4 (Figure 1) which due to its age and other challenges such as increase in energy consumption (comparable to a city with 10,000 inhabitants) will be replaced by the VSC 5 (Figure 2) which is more efficient. The system is cooled down using water and does this 'greenly' by not depending on cooling devices such as ACs because of the pumping system that sends the heated water to the overhead water tanks where the waste heat dissipates directly to the atmosphere.

Additionally, we also toured the Tape room (Figure 4) which is where the archive data is stored and is only accessible via request. It operates such that when a request is made, the robot arm (see Figure 5) inside the storage picks the magnetic drive and inserts it into the reader where the user can now access the data requested.

The lecture session also introduced some future advancements the institution is involved and during the guided tour we also saw The Multi-Site Computer Austria (**MUSICA**), Austria's next supercomputer, which is a collaboration between several universities in Vienna, Innsbruck, and Linz. The goal is an expansion of the VSC where the computer hardware will be distributed across several locations, combining HPC with cloud computing significantly increasing the computing power for users. Some challenges that this initiative poses include setting up of all the locations infrastructure, data transfers, management, and standardization across all the hubs.

Moreover building on the VSC, the development of AI factory Austria (AI:AT) a new AI-optimized supercomputer is also underway and will be available to researchers, startups, companies, and public administration to enable complex AI models to be efficiently trained and transferred to various application areas.

Reflection and Conclusion

With the vast increase in EO data, EODC plays an important role in the storage, distribution, and processing of this data. This excursion fulfilled both the course and my personal objectives and in addition I was able to see an actual supercomputer (aside from in the movies) understand it components and importantly its challenges especially cooling and energy efficiency.

The excursion also raised some questions for me such as: why not locate the supercomputers in cooler locations like the Austrian Alps have (aside from Antarctica, as was suggested during the excursion) and find a solution for the latency in internet connection if that arises, given the Alps is closer in proximity compared to Antarctica, because of how hot it gets in Vienna especially during the summers reducing the efficiency of the machine.

Reflecting to my country of origin, Kenya, I am curious about how feasible it is implementing a supercomputer. Its not lost on me that it is expensive to set up a supercomputer but nonetheless this excursion, raised some questions to this regard: What would be the uptake of this by users? would the government/politics be for or against such a set up? Do we set up a supercomputer for the continent or at least the region (Eastern Africa) if not feasible for the country alone? I am inclined to think that the lack of good collaboration channels among the universities within Kenya, as it is in Austria and Europe, will make it harder to set up this infrastructure in a similar manner. The implementation of a supercomputer is something worth exploring/implementing in my country/region especially also given the existence of data cubes such as Digital Earth Africa generating even more EO data would hopefully lead to more effort in the aspect of setting up these computing resources.

Some pictures I took during the guided tour:

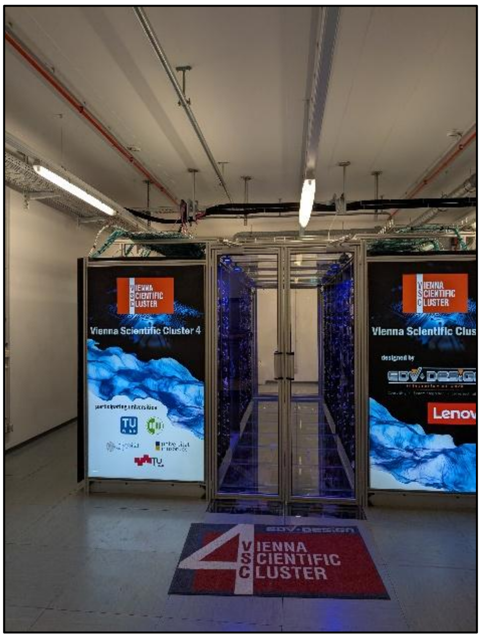

Figure 1: VSC 4

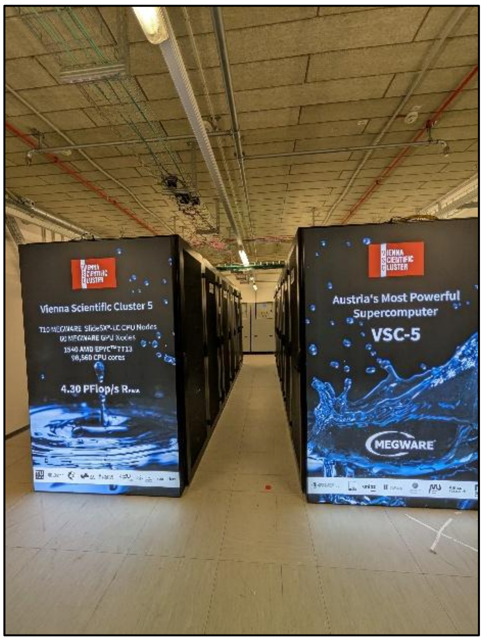

Figure 2: VSC 5

Figure 3 MUSICA

Figure 4 Tape room

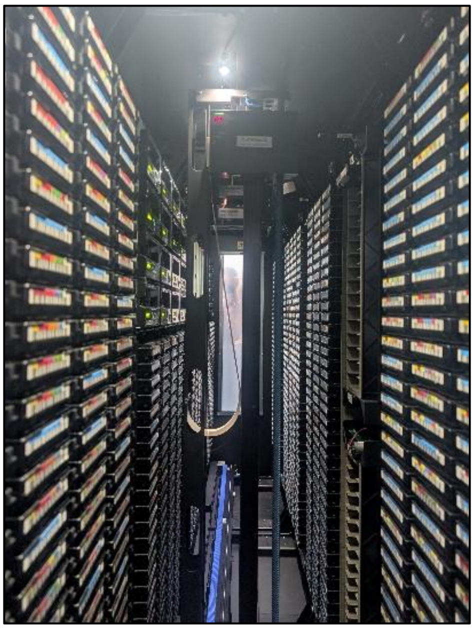

Figure 3 Robot arm in the tape room

Environment Agency Austria

Introduction

The emergence of Big Data in EO has been increased majorly due to the readily available datasets produced by various sources making data users more aware of the value obtained from combining and linking different data sets produced by different producers, public or private (Open Government Data, 2013). Therefore, there is the necessity to ensure the variety of this data is open and accessible for users, government included, to perform their tasks. This where the concept of open government data comes into play. Open government data is defined as any data or information which is produced or commissioned by public bodies, can be freely used, re-used and distributed by anyone provided users properly attribute it for sharing (Open Government Data, 2013). This enables improved service delivery, transparency, and public engagement and as a result better relations between governments and citizens.

Overview of UBV

The Environment agency Austria (Umweltbundesamt - UBV) is one such institution that produces or commissions open government data. It is a government owned agency that provides important decision-making bases for proactively addressing challenges at the interface of the environment, economy, and society. It includes the departments that are responsible for land cover, biodiversity and water, climate, habitat networking, GIS and biodiversity monitoring.

During the visit, at UBV offices in Vienna, we had a session led by **Dipl.-Ing. Gebhard Banko and Roland Grillmayer** who gave us a comprehensive overview of UBV's operations and projects spanning various domains within remote sensing and spatial analysis and the role that the environment agency of Austria plays. The resulting data and tools of the projects are open government data that abide by the FAIR principles in that they are Findable, Accessible, Interoperable, and Reusable.

One of UBV's project domains is in remote sensing applications include monitoring land take and soil sealing by combining administrative GIS data with EO data to enable operational tracking of settlement areas, traffic and leisure activities. It also supports *nature restoration* by mapping green urban areas and tree crown cover in urban ecosystems as well as assessing vegetation dynamics across various spatial temporal scales. In biodiversity, UBV designs and maintains data sets and visualizes the important habitat corridors in Austria. It also assesses the impact of infrastructure in the protection of biodiversity. UBV, being a key partner institution to Copernicus, has a some Copernicus projects including assessing the suitability of Copernicus portfolio, monitoring inland water, STAC for Copernicus land monitoring service products on CDSE , *CORINE land cover 2024 – update* .

Moreover, the agency is involved in developing tools that can be utilized by the public, government, and private stakeholders for various use cases. These tools include DROP where UBV develops machine learning algorithms to map irrigated fields over four intensively irrigated Austrian regions; **REST-GDI-AGRAR** a toolbox for the implementation of EU common spaces; *Environmental data for the semantic web and AI* for the provision of geodata in Knowledge Graphs and the Semantic Web and Digital twins based on knowledge graphs for complex time series analyses. Additionally, the agency offers knowledge transfer through capacity building projects. These projects funded by the EU include *EU4Environment – Water and Data* and EU4Green Recovery East

Conclusion and Reflection

EO data is produced in vast amounts form both public and private institutions. Having systems in place to ensure this data is findable, accessible, interoperable, and reusable is important. This will reduce redundancy in data, ensure efficient data storage, transfer between users and stakeholders and data maintenance. Government plays a key role in putting in place policies or regulations that ensures these practices are upheld.

The excursion provided a clear picture of the operations and projects the environment agency Austria is involved in and how the link back to Copernicus data space ecosystem. The showcase of the tools they have developed for data dissemination and implementation of use cases using unique programming i.e. knowledge graphs and Neo4js which utilizes graph SQL for semantic querying was an eye opener to different tools/software in use in the EO domain. Their utilization of open geospatial standards in setting up their web tools was also showcased and was impressive to see as this concept was tackled in one my electives course (IP: SDI) where we were tasked with developing an SDI architecture. Seeing the GDI they developed cemented the importance of open standards especially for government data.

Lastly, reflecting to my county and the continent at large with regards to the policy of open government data as is established in the EU and Austria. Kenya for example produces a lot of data that are GI related however standardization is lacking especially regarding naming. For example, in my previous role in Kenya, I noticed there is little to no naming convention for regions, or rather not enforced, for the different agencies that produce maps with regional boundaries. This made seeing how the environment agency of Austria has set up the GDI to ensure semantic querying considered even naming that differ region-wise very impressive. This feature proves to be very essential in all data infrastructures globally because it minimizes redundancy and ensures efficiency due to easy access for users and data providers as well as enabling government accountability.

EU Inspire & EU Copernicus Data Space Ecosystem

The EU has set up initiatives to handle the increasing amount of EO data for different use cases. Two main initiatives include the EU Inspire directive and EU Copernicus data space ecosystem.

EU Inspire

The EU Infrastructure for Spatial Information in the European Community (INSPIRE) came into effect in 2017 as a directive whose main goal is to establish a spatial infrastructure for sharing of environmental spatial information among public sector organizations and better facilitate public access to spatial information across Europe. Its establishment aimed to solve several challenges among the member states of the EU. These include: a variety of data formats, different data semantics, differences in data quality, poor documentation, gaps in geographical coverage and incompatible systems(*Inspire_introduction_sudra*, 2025.).

INSPIRE involves five key components: metadata, data specifications, network services, data and service sharing, and monitoring and reporting, with detailed implementing rules(Bray & Ramage, 2012.). The directive is based on the following principles: data should be collected once and maintained effectively, spatial data should be interoperable in that combined easily and shared between many users and applications, should be readily available and transparent, and should be properly documented making it easy to find what information is available, and under which conditions it can be acquired and used(*Inspire_introduction_sudra*, 2025.).

The directive has 34 spatial data themes needed for environmental applications divided into 3 Annexes(Bray & Ramage, 2012.). To achieve its goals, the directive relies on the OGC standards. The key open standards outlined in the framework include Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS), Web Map Tile Service (WMTS), Catalogue Service for the Web (CSW) and Geography Markup Language (GML).

EU Copernicus Data Space Ecosystem

The Copernicus Data Space Ecosystem (CDSE) was launched in 2023 as an open ecosystem that provides free instant access to a wide range of data and services from the Copernicus Sentinel missions, including Sentinel-1, -2, -3, -5P, and related processing capabilities. This enables the development of value-added services such as the six thematic Copernicus services: Land Monitoring, Marine Environment Monitoring, Atmosphere Monitoring, Climate Change, Security, and Emergency Management.

The ecosystem offers a range of services, infrastructure and tools for users to efficiently access and process sentinel data, integrate their own datasets and deploy applications at various levels of federation. The key components include data (Copernicus Sentinels Missions, Copernicus Contributing Missions, Federated data sets), data workspaces (for cloud computing i.e. Jupyter lab environments), APIs for data access and processing capabilities (including S3, STAC, SentinelHub APIs) and applications such as the Copernicus browser for data search, access and visualization.

REFERENCES

Backeberg, B., Šustr, Z., Fernández, E., Donchyts, G., Haag, A., Oonk, J. B. R., Venekamp, G., Schumacher, B., Reimond, S., & Chatzikyriakou, C. (2023). An open compute and data federation as an alternative to monolithic infrastructures for big Earth data analytics. *Big Earth Data*, *7*(3), 812–830. https://doi.org/10.1080/20964471.2022.2094953

Bray, A., & Ramage, S. (2012). *OGC Market Report Open Standards and INSPIRE*.

European Commission. Joint Research Centre. (2023). *Proceedings of the 2023 conference on Big Data from Space (BiDS'23): From foresight to impact : 6 9 November 2023, Austrian Center, Vienna.* Publications Office. https://data.europa.eu/doi/10.2760/46796

Inspire_introduction_sudra. Online access (2025).*INSPIRE Directive.*

Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives (OECD Working Papers on Public Governance No. 22; OECD Working Papers on Public Governance, Vol. 22). (2013). https://doi.org/10.1787/5k46bj4f03s7-en

Sudmanns, M., Augustin, H., Killough, B., Giuliani, G., Tiede, D., Leith, A., Yuan, F., & Lewis, A. (2023). Think global, cube local: An Earth Observation Data Cube's contribution to the Digital Earth vision. *Big Earth Data*, *7*(3), 831–859. https://doi.org/10.1080/20964471.2022.2099236

Wagner, W., Fröhlich, J., Wotawa, G., Stowasser, R., Staudinger, M., Hoffmann, C., Walli, A., Federspiel, C., Aspetsberger, M., Atzberger, C., Briese, C., Notarnicola, C., Zebisch, M., Boresch, A., Enenkel, M., Kidd, R., Von Beringe, A., Hasenauer, S., Naeimi, V., & Mücke, W. (2014). Addressing Grand Challenges in Earth Observation Science: The Earth Observation Data Centre for Water Resources Monitoring. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *II–7*, 81–88. https://doi.org/10.5194/isprsannals-II-7-81-2014