DEGREE PROJECT IN MEDIA TECHNOLOGY,
SECOND CYCLE, 30 CREDITS
STOCKHOLM, SWEDEN 2018

Visualization and Interaction with Temporal Data using Data Cubes in the Global Earth Observation System of Systems

JOAKIM ADRUP

KTH ROYAL INSTITUTE OF TECHNOLOGY
SCHOOL OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

Visualization and Interaction with Temporal Data using Data Cubes in the Global Earth Observation System of Systems

Visualisering och Interaktion av Tidsbaserad Data genom användning av Data Cubes inom Global Earth Observation System of Systems

Joakim Adrup
Royal Institute of Technology
Stockholm, Sweden
ESA's Centre for Earth Observation
Frascati, Italy
joakim@adrup.com

ABSTRACT - ENGLISH

The purpose of this study was to explore the usage of data cubes in the context of the Global Earth Observation System of Systems (GEOSS). This study investigated what added benefit could be provided to users of the GEOSS platform by utilizing the capabilities of data cubes. Data cubes in earth observation is a concept for how data should be handled and provided by a data server. It includes aspects such as flexible extraction of subsets and processing capabilities. In this study it was found that the most frequent use case for data cubes was time analysis. One of the main services provided by the GEOSS portal was the discovery and inspection of datasets. In the study a timeline interface was constructed to facilitate the exploration and inspection of datasets with a temporal dimension. The datasets were provided by a data cube, and made use of the data cubes capabilities in retrieving subsets of data along any arbitrary axis. A usability evaluation was conducted on the timeline interface to gain insight into the users requirements and user satisfaction. The results showed that the design worked well in many regards, ranking high in user satisfaction. On a number of points the study highlighted areas of improvement. Providing insight into important design limitations and challenges together with suggestions on how these could be approached in different ways.

ABSTRACT - SVENSKA

Syftet med studien var att undersöka hur Data Cubes kunde komma att användas inom ramarna för Global Earth Observation System of Systems (GEOSS). Vilka fördelar som kunde dras ifrån att utnyttja den potential som data cubes besitter och använda dem i GEOSS plattformen undersöktes i studien. Data cubes för earth observation är ett koncept om hur data ska hanteras och tillhandahållas av datatjänster. Det ämnar bland annat flexibel extrahering av datapartitioner och dataprocesseringsförmågor. I denna studie iakttogs det att det mest frekvent förekommande användningsområdet för data cubes var analys av tid. Ett huvudsyfte med GEOSS portalen var att tillhandahålla användaren med verktyg för att utforska och inspektera dataset. I denna studie tillverkades ett användargränssnitt med en tidslinje för att ge användaren tillgång till att även utforska och inspektera dataset med en tidsdimension. Datasetet tillhandahålls från en data cube och utnyttjar data cubes färdighet i att förse utvalda partitioner av datasetet som kan extraheras längs valfri axel. En användarstudie har gjorts på användargränssnittet för att utvärdera till vilken grad användarna var nöjda och hur det uppfyllde deras krav, för att samla värdefulla insikter. Resultatet visar på att designen presterar väl på flera punkter, den rankar högt i användartillfredsställelse. Med studien klargör även framtida förbättringsmöjligheter och gav insikter om viktiga designbegränsningar och utmaningar. I rapporten diskuteras det hur dessa kan hanteras på olika sätt.

KEYWORDS

visualization, data cube, GEOSS, temporal data, geographical, usability, user satisfaction, affordance, limitations

1 INTRODUCTION

Data cubes provide a range of different capabilities. Which makes them powerful tools, but creating intuitive interaction with them through a user interface is not trivial. This study is done in order to identify ways that data cubes can be utilized in the framework of the Global Earth Observation System of Systems (GEOSS) by gathering insights into which approaches are the most beneficial to communities interested in Earth Observation. One of the main goals for GEOSS is to provide all manner of earth observation products to world wide communities [2]. Lately there has been an increase in adoption and promotion of data cube technology [4, 13, 17]. The result is an untaped source of earth observation data that the GEOSS platform ideally should also be able to provide to its users. Data cubes comes with additional benefits that could be harnessed inside the GEOSS platform and portal for the benefit of its users.

This study is a continuation of the work presented at PV2018 under the title "Pioneering Steps towards Use of Data-cubes in the Global Earth Observation System of Systems" [1]. A prototype demo was also presented at the 3rd GEO Data Providers Workshop held at the European Space Agency in Frascati, Italy. The prototype has since then undergone refinements and is presented here in the method section. This report provides more extensive background information as well as user studies and insights gained from these.

1.1 Earth Observation Data Cubes

Data cube is a concept for multidimensional data structures. In the field of Earth Observation (EO) there has been an initiative for a common format and concept of what a data cube should entail [28]. The aim of the concept is to have a unified view of what tools should be available to users of data cubes. Some of the defining characteristics and features of EO data cubes that set them apart from other multidimensional data structures are: they should be implemented in a way that supports slicing and trimming in all axes, data axes should be treated the same way no matter the data represented, data cubes should perform equally well in extracting data in any axes [7]. As concisely explained and summarized by the data cube manifesto:

> " ..it has the potential of greatly simplifying Big Earth Data services for users by providing massive spatio-temporal data in an analysis-ready way [7]".

Analysis ready data (ARD) is provided for the convenience of the data scientists and other users. The data is preprocessed into products that can be downloaded and used in analysis without further effort and thus, it relieves the user of the time consuming task that comes with data preparation. Even though data cubes seem to imply three dimensions, it is shorthand for hyper-cubes and supports arbitrarily many dimensions. It is important to note that it is also meant to handle fewer dimensions than three. What the axes represent depend on the associated meta data. Which means that the raw data can be handled exactly the same for any dimension. This feature allows for a more concise and easy to work with toolset (figure 1).

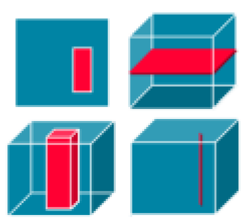


Figure 1: datacube: sample overview of how subsets can be extracted [7]

The data cube used in this study is called EO Data Cube and is a product of the European space agency's EO Datacube initiative. The data cube contains Sentinel 1,2, Landsat 7,8 and various other products such as MODIS surface temperature and NDVI. The data is provided as a Web Coverage Service (WCS). See terminology section for description of acronyms and satellite specific terms.

There is an initiative called open data cube that aims to make the access to data cube technology publicly available as open source in an effort to promote a more widespread use [7, 28]. There are already many successful implementations of data cubes. One is called the Australian Geoscience Data Cube or AGDC [17]. It is a collaboration of multiple different parties in Australia that holds an interest in EO data. One of their purposes for which they utilized the cube is to monitor flooding river valleys. Since the Landsat program has been operational for over 40 years there is a wealth of historical data to analyze as well [31], giving opportunity to investigate previous floods and how they affected the environment.

Another successful implementation is the Swiss data cube [13]. The purpose of the Swiss data cube is to provide Landsat data, making it easily available for analysis in order to tackle growing resource, economical and social demands. To make the list more complete there is also the Colombian (CDCOL), Committee on Earth Observation Satellites (CEOS), Land Change Monitoring Assessment and Projection (LCMAP) and The Commonwealth Scientific and Industrial Research Organization (CSIRO) who all implement their own data cubes [4, 16, 29]. It is also important to know that the concept of data cubes and especially multidimensional data structures has a long history. Data cubes exist in different form and applications, wherever multidimensional data structures are useful. This text only focuses on Earth observation data cubes and the push for a common definition of a data cube specifically adapted for this domain of work and research.

1.2 GEOSS platform and portal

> "The GEOSS Portal is one of the world's few global systems for accessing EO data, as well as government, organizational and private sector data for remote sensing, in situ and atmospheric data. [2]"

Global Earth Observation System of Systems (GEOSS) is an initiative to facilitate and promote the usage of EO data. Earth observations are a useful tool for understanding future and past climate change. Making EO data easily discoverable and available is essential to help in dealing with future challenges on disaster resilience, agriculture, rising water, etc [21]. The GEOSS initiative provides the GEOSS portal as well as the GEOSS platform. The GEOSS portal is a web interface that provides users with the capabilities to search for EO data, view the datasets on a map interface and then download the data. The GEOSS platform architecture is a modular approach to the different services that GEOSS provides [30]. One part of the GEOSS platform is the GEO Data Access Broker (DAB) [21, 22]. The broker is the middleware of the GEOSS platform and portal. It is tasked with keeping track of all the contributing sources of data [8, 9, 22]. It allows users to connect and find data. DAB handles metadata that links to the actual data repositories. Metadata is all the data that describes the dataset. Information such as location, acquisition time, format, organization info, etc. DAB handles converting between many different input formats and translates these into one output format. Making the diverse set of data more accessible. DAB can be accessed by the GEOSS portal and through various other resources. A REST API and a JavaScript library is provided. Users can create database views that contain only the information that is relevant to them. Making the task of narrowing down the data search result significantly easier.

Aside from that users can implement the whole GEOSS portal search box in their own gateway as a widget, which supplies the search functionality without much development on the user side. Mirror sites of the GEOSS portal can be created upon request. These are then combined with data views suited to the users needs. Part of the GEOSS platform is the Yellow Pages. The Yellow Pages is designed as an entry point for new data providers as well as giving information on current contributors. Lastly the platform also includes the status checker, which is a service that monitors the health of the data providers repositories. It is used for results ranking, providing higher rankings to datasets that are more reliable.

GEOSS currently has more than 100 member countries that contribute as data providers and data users. Data is not limited to satellite data but can be anything EO related [9]. A provider can supply data for example from weather stations. Near earth or field observations are usually called in-situ observation. Since EO data does not have to come from satellites that means anyone collecting data can potentially be a data provider [2]. One of the challenges that the GEOSS platform faces is the overwhelming amount of data, data that can be extremely valuable if harnessed for analysis and decision making [12, 22]. This type of resource is commonly refered to as Big data. Big Data is usually defined through the three V's: volume, velocity and variety [12]. Volume simply refers to the magnitude of data, velocity is the frequency at which the data is generated and processed, and variety is to what degree of disorder the data is in. The term is often extended with additional defining terms, Veracity, Value and Visualization. Veracity is the degree of unreliability of the data and value is the benefit that can be gained from the data. In big data the value is often low per byte and the task is to refine the value into something meaningful [12]. Visualization refers to the challenges of creating insightful visualizations of the big data. The GEOSS platform and especially the GEO DAB deals with big data in every remark according to this definition [22].

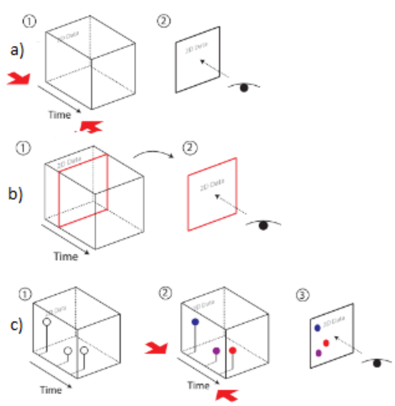


Figure 2: Time operations: sample overview of how time can be viewed in 2D [6] a. Time flattening, b. Time slicing, c. Colored time flattening

1.3 Visualize the data cube

How do you visualize the data cube? One method is to cut a time slice. Which means extracting one frame of raster data at a specific time point [6]. This method is the most straightforward. It is also a common concept, retrieving a frame from a video (see figure 2). This concept can generally always be applied no matter the dataset representation. Another approach is called time flattening. You do this by combining multiple different time points into one picture. This can be extended with colored time flattening, where each time point is assigned a specific color. Which area of the image that is affected by a certain time point can then be derived from the colors found in the image [6].

The common way to represent time is by a time axis. You can then animate the single time points into a series, effectively creating a movie. Resch argues that it should be investigated if this really is the best way to understand a temporal dimension [26]. One method of approach is to use the spacial z-axis to instead signify temporal data [3]. This can work well when the significant temporal events are sparse, the limitation is that it becomes cluttered for easy reading when the data is not sparse enough.

1.4 Designing interaction

The design of visual cues and the interaction plays a large role in the usability of the product [32]. In Human-Computer Interaction (HCI) and User Experience (UX) the concept of affordance can be interesting to evaluate and keep in mind while designing for the user. The concept is defined by the following attributes:

- Suggestions or clues as to how to use the properties.
- Can be dependent on the experience, knowledge, or culture of the actor.
- Can make an action difficult or easy.
- Perceived properties that may or may not actually exist.

This interpretation is defined by Norman and concisely described here by McGrenere and Ho [20, 24]. It is good practice to consider how an interactive element is seen by the user. Does it visually convey how it should be used. For example a button that is raised gives the impression that is can be pushed. Which logically leads the user to think that it will trigger some type of functionality. A slider has the affordance that it can be pushed along the visual axis, usually along the screen x or y-axis. It is important that the users anticipated outcome from the action relates to the actual outcome. When designing it is important to understand which symbols and icons are the most logical choice for the given operation [27]. An operation could be to send a web form or exit and application. The icons should be clear and accurate, easily and correctly interpretable and should give visual feedback when in use. The interface should also make information available on how the tool is supposed to be used, e. g. by giving a tool tip while hovering. For example explaining why a tool is not available at the current state of the website.

1.5 Research Question

What are the affordances and limitations of a graphical interface between the GEOSS platform and the EO Data Cube for the purpose of inspecting and finding items in geographical datasets with a temporal dimension aimed for a community familiar with earth observation and measured by surveys and interviews investigating perceived user satisfaction?

2 RELATED WORK

There have been multiple attempts to create a visualization of the time axis in a more intuitive way. It is very beneficial to see all the information in one glance. But it is always a balance between providing not enough information or too much. Resch attempts to visualize lightning strikes as an animation of lightning visuals striking a terrain map [26]. Resch compared this to a heat map that showed the same type of information. The study had some success but the result was divided. There has also been some attempts to visualize the time axis as the spatial z-axis. The height of the event shows at what point in time it occurred [3].

There is a lot of UX research on the topic of sliders. Sliders in graphical user interfaces generally intend to give the user a predefined range of options. This allows the user freedom while also clearly conveying the limitations for the user. The slider also has the affordance that it looks like either a horizontal or vertical axis. It is intuitive to allow the user to choose a value that could be lined up along that axis. So it is a natural choice for representing a dimensional axis. Laubheimer argues that sliders can be good for exploration, but the developer should be careful about the balance between exploration and precision [15]. With more flexibility on the range of options the more frustrating it can become for a user that wants to choose a specific value. Babich argues that the metaphor for sliders work well when the data can be perceived as a continuum. Such as the volume for a loudspeaker or the progression of time in a video [5].

Given the difficulty in visualizing geographical data with a temporal dimension, as discussed by Resch [26]. I set out to investigate graphical interfaces that allows the user an intuitive visual understanding of the data extent contained in the temporal dimension of data provided by a data cube. The interface should also allow the user to interactively inspect the contained content.

3 TERMINOLOGY

3.1 Analysis Ready Data

Analysis ready data (ARD) refers to data that has be processed from the initial captured data into a format that is easier for a scientist or analyst to use right away. The process to get to this stage always depends on the product. What product is relevant depends on what analysis the user is interested in performing. Different capturing techniques require different processes. One example of this is that optical images suffer occluded sight under poor cloud conditions. Clouds and cloud shadows are often removed from the data. However radars do not suffer the same problem since radio waves can pass through clouds.

3.2 Sentinel

Sentinel is a series of satellite missions developed by the European Space Agency. Their purpose is earth observations for monitoring of environment and security issues, such as disaster control and prevention.

3.3 Landsat

Landsat is a series of Earth observation satellites managed by NASA. They provide information about various climate aspects, such as carbondioxide, land cover, etc.

3.4 MODIS

Stands for Moderate Resolution Imaging Spectroradiometer. It is a satellite instrument provided by NASA that has a range of different products, e. g. land cover, radiation, reflectance, NDVI, etc.

3.5 NDVI

Normalized difference vegetation index. Method for measuring and visualizing the degree of vegetation in an area.

3.6 Tiles and tile service

With sizeable images memory and performance becomes an issue. A more sophisticated and commonly applied method is to prepare the image into a tile pyramid. The vertical layers of the pyramid relates to the users zoom level. Each layer is partitioned in a grid of tiles, each having their respective x and y position on the grid. This enables the application to only load the required amount of pixels. The viewing application can omit tiles outside of the screens view window. The application can also supply a lower resolution image when the full resolution is not required, in other words if the user is viewing the image from afar. Tile services provide an interface for retrieving tiles, usually by specifying a tile with coordinates such as x, y and z. A common format is called Web Map Tile Service (WMTS) and is defined by the Open Geospatial Consortium (OGC).

3.7 Geospatial web protocols

Protocol mentioned in this paper are the following WMS, WMTS, WCS and WCPS. All of them are defined by the Open Geospatial Consortium (OGC) and can be found via their website opengeospatial.org. Web Map Service (WMS) is a common way of providing two dimensional geo-referenced imagery. It can be extended with time and elevation parameters that designate which time or elevation slice that is intended for retrieval. WMTS provides maps but as tiles (See section about Tiles and tile services). Web Coverage Service (WCS) offers more flexibility for providing geospatial data with multiple dimensions. WCPS is a processing protocol usually used in combination with WCS. Allowing users access to server side processing of the data before it is provided.

4 METHOD

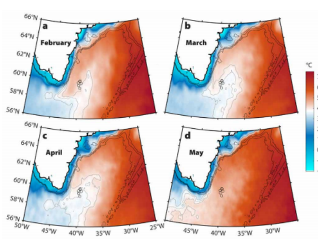
This section will explain how preliminary work served to justify the investigation into a time oriented interfaces and on what grounds a prototype was conceived and constructed. It also entails the method of inquiring about user satisfaction as a measure of prototype performance.

4.1 User Scenarios

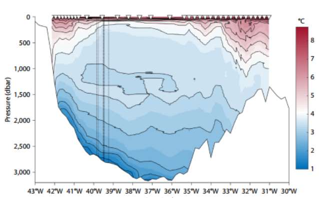
As preliminary work we studied communities, these were picked because they could have an interest in the GEOSS platform and earth observation data. These communities have the common denominator that they all have an interest in climate and environment in one way or another. However it is not apparent how data cube capabilities is best applied in order to support these activities and therefore motivates the method of defining user scenarios for each community. A community is deemed to have an interest in GEOSS either from previous collaboration with the organization or by actively using data that can be discovered in the GEOSS portal and platform. A simple approach for researching communities was used. Official material was examined such as websites and publications. In this material the important information extracted were the following: the goal of the organization and what parameters they were interested in, for example temperature, land cover, deforestation, etc. One type of EO data used by these communities were satellite products such as optical, radar, near-infrared, infrared, surface reflectance, etc. The products are commonly stored as two dimensional raster images and often combine multiple color bands. The products also have information about acquisition date and time. Some datasets are already processed into level 4 products, level 4 signifies that they are derived from lower level products such as raw satellite data. An example of this is how the Global Forest Observation Initiative (GFOI) utilizes forest/non-forest maps provided by the Japanese Space Exploration Agency (JAXA) [11].

There are also examples where communities have had use for EO products with a vertical dimension. This dimension were especially prevalent in oceanography, where understanding what happened in the interior of the ocean was important [10]. AtlantOS and International Ocean Carbon Coordination Project (IOCCP) have shown interest in this type of data. The data is collected in-situ, via buoys or ships.

The most common type of visualization of geographical data occurring in publications from communities were: two dimensional raster images, graphs and scatter plots for correlations between variables. The two dimensional raster images are most often colored with a gradient that illustrate the important areas of the visualization. It has been found that these types of visualization appear in combination with each other. Sometimes vertical slices are also used when three dimensional data is relevant. In figure 3, two images have been retrieved as examples of how data can be presented. Both were taken from the same report which illustrates how both methods can successfully co-exist in a publication in order to enhance the message [10].



(a) Latitude and Longitude slice

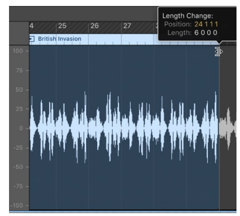
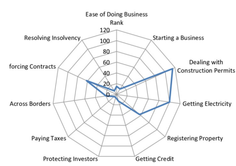
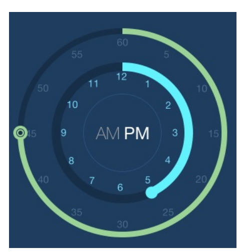


(b) Vertical slice

Figure 3: Examples of how map imagery and vertical slices can be visualized in publications [10]

The time aspect was ubiquitous in one way or another for every community. This inspired the notion that a tool for browsing time conveniently, would be one of the most essential additions that data cubes could provide to users. Therefore it became the starting

Table 1: Different techniques for viewing and selecting time related content

Sample view	Advantages and Disadvantages
	Animation editing (Photoshop) - Quick preview of each frame. - Play capabilities. - Playback speed settings. - Becomes cluttered easily with multiple frames. - Timesteps are meassured in number of frames. - Extra latency on system to generate thumbnails.
	Audio editing (GarageBand) - Easy to select a range. - Intuitive overview of existing data. - Hard to select a single sample. - Diminishing usefulness of data overview when zoomed out. - Measurement units is samples over time.
	Spider chart - Allows for arbitrary dimensions. - Occupies a lot of screen space. - Non intuitive what each slider step signifies. - Difficult to have precise selection with high number of steps.
	Circular slider - Gives an intuitive analogy to time due to the clock like visuals. - Commonly seen in applications. - Works well on a touch interface. - Easy to select region. - Occupies a lot of screen space. - Hard to select region spanning more than a year.
	TimelineJs - Intuitive overview over large time intervals. - Accommodates a diverse set of content, videos, text, images, etc. - Can visualize time ranges. - Limited selection capabilities for ranges. - Only works well for sparse information.

point for this paper. For the full list of investigated organizations see appendix.

4.2 Interactive techniques for time related content

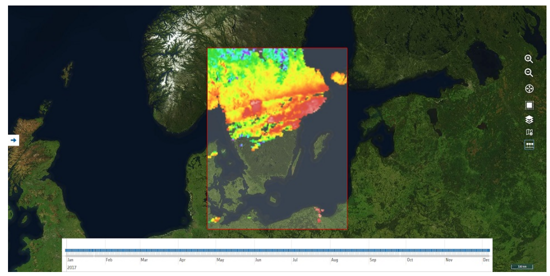
The technique used for manipulating and showing time related content was inspired by TimelineJs (see table 1). To counter the problem of not supporting dense information sets the text and icon content was kept to a minimum. A clustering algorithm was also introduced that grouped items when zooming. The feature of selecting ranges was omitted, instead precise selection of discrete items was favored. The reason was that precise date picking would be useful for the user and comparing dates would be easier.

4.3 Prototype

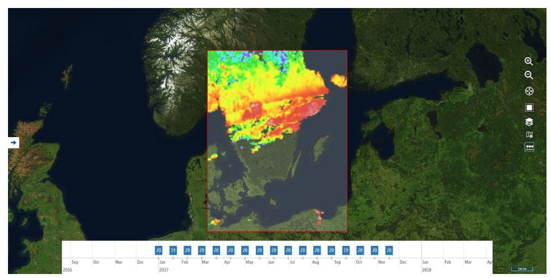
The prototype was based upon the geoportal.org, as it looked like in April of 2018 [25]. The vis.js timeline module was used for the visualization. In the tool bar the user could select an area of interest and then open the timeline tool to view the available data points (see figure 4a). The points were shown as markers on the dates that they were acquired. Showing many data points comes with the problem that they could easily become crowded and overlap each other. This led to the implementation of a clustering algorithm (see figure 4c). Clustering in this context is the act of grouping object together when you zoom out and view a larger range. It is naturally complemented with an expansion of the clusters when the user is closer to the individual items (see figure 4d). The relative distance in seconds between objects is compared to the whole visible time frame. If the distance is comparatively small then items are clustered and vice versa. Another common approach is to stack items above each other when they collide. This works reasonably well when 2 - 3 items collide. However in this prototype the items usually number in the range 200 - 300 which renders that method entirely inconvenient.

The time-line supports panning left and right by clicking and holding the mouse button inside the timeline tool and then do a dragging gesture either left or right. It also supports zooming in and out, making the task of browsing dates over a large time ranges easier. When the user clicks on a date the data corresponding to that date is fetched from the server and display inside the area of interest. The extent of the data requested is cropped to only include the range defined by the area of interest, one of the benefits provided by data cubes. This makes the website more responsive since less data has to be loaded. The active date is highlighted in a light green. The highlight transfers to the cluster that consumes the item while scrolling, allowing the user to always know roughly where the selection is.

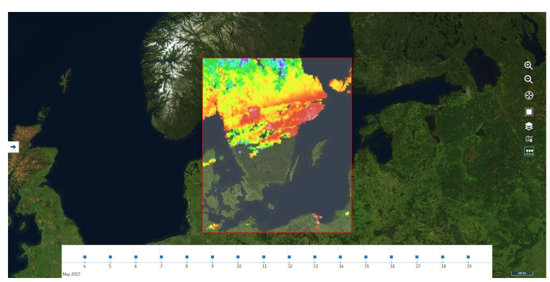
(a) Selecting area of interest.



(b) Data loaded by using timeline.



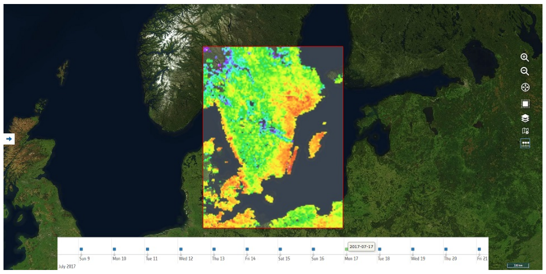
(c) Clustering of timeline objects when zooming out on the timeline bar.



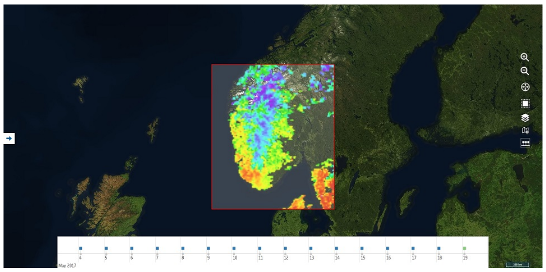
(d) Declustering of timeline objects when zooming in on the timeline bar.

4.4 User Study

The purpose of the study was to test the prototype for how well it conveys information about data points in a data cube as well as facilitating dynamic interaction with the data points. The focus is for the user to easily find specific information that they are looking for. Such as dates in the timeline or features of the data for example where it is missing or incomplete. How well the prototype performed was measured by user satisfaction accompanied by inquiring interview questions. Inspecting data can take as much time as the user is willing to spend on it. The amount of detail the user chooses to go into can vary a lot, which makes it non nonsensical to measure the efficiency of the prototype. Instead it was highly important that when the user wanted to do something, the prototype was not limiting the person. I chose to investigate user satisfaction



(e) Change of time point in the timeline.



(f) Change of area of interest.

Figure 4: Timeline tool, example data is MODIS surface temperature.

as a measure on how limiting or enabling the prototype was to the user.

The participants were recruited from the ESA staff and contractors present in the ESRIN area. The choice of users already familiar with EO was convenient in order to identify problems that scientists could have. Participants had seen similar applications before and could give valuable inputs on how the prototype was to evolve in the future to fit their and others needs better. It was also the case of proximity and convenience.

Tasks were given to the users upon arrival. The tasks included locating the timeline tool in the portal, loading data for an area, looking for dates and browsing nearby dates. The task of browsing nearby dates was meant to simulate looking for features in the dataset. The feature that was used in this study was simply to find a complete coverage of the date selected. The data that was used had some dates with lacking coverage. Other scenarios of browsing for features could easily be thought of. One could for example be to identify local heat waves in temperature data. Another purpose of having the user look for dates was that it worked to test the capabilities of the timeline. Testing how zooming, panning, selecting, clustering, etc, worked for the user. Giving them incentive to explore all available interactions.

A questionnaire was created to help evaluate the prototype. It was made using the UMUX-Lite way of inquiring about user satisfaction [19]. More precise user input was required however. So it was extended to inquire about specific elements in the design as well. These questions also followed the UMUX-Lite way of asking questions and grading answers. The questions gave a statement such as "The time-line capabilities meet my requirements" that the participant could grade how true the statement was on a scale 1 - 7, where 1 meant that the user "disagrees fully" and 7 meant "agrees fully".

An interview was also conducted to complement the questionnaire. The interview was designed to give the participants more room for complex feedback, such as what elements worked, what elements were less effective, which elements were perceived to have a different function compared to what the user initially thought. These questions helped pinpoint design errors that were not anticipated and design opportunities that could be further considered. It also gave the user room to discuss the reasoning they had for statements they made while doing the test and answering the questionnaire. The cause behind problems are really the essence of figuring out limitations in the design and affordances. This process is essential but tricky to make consistent. When a user gives an explanation it is often interesting to follow up with more detailed questions related to the explanation. Causing each interview to be very unique but also very informative.

Observations were carried out while the participants were using the prototype. It helped identify how the user approaches the prototype. It also gives hints to what elements are unclear. It gives insight into events that the user themselves do not tell you about. Many smaller things might seem trivial to the user which makes them forgo telling the tester about them. Observing allows for catching these smaller aspects that might in the end be crucial.

4.5 Extended User Study

During the 3rd GEO Data Providers Workshop a hackathon was conducted. One of the purposes was to investigate the usability of the GEOSS platform and data access requirements. A survey was handed to the participants afterwards. A handful of the questions were designed to inquiry about the perceived relevance for accessing time related data, interacting with time related data, comparing datasets and comparing time ranges. It was done in order to gain understanding if the project was on the right course and to further cement the ideas hatched in the user scenario study. Questions about the perceived relevance for different data cube functionalities were also included in the study and paper.

5 RESULT

The study was conducted on 8 participants. Of which 2 were female and 6 male. Participants were all in the age group 25 - 34 years old. All participants were either staff, contractors or trainees at ESA's Centre for Earth Observation. They all had experience and familiarity with earth observation data and the navigation of map interfaces. They are considered expert users in this work context and with this type of interfaces. In usability studies the the common rule of thumb is that 5 participants usually find about 80% of the flaws in a design [18]. The sample size of 8 in this study is therefore deemed to be sufficient. There was also a survey handed out to the participants of the GEOSSHACK that ran in parallel with the 3rd GEO Data Providers Workshop. 11 people gave responses in that survey.

5.1 Quantitative User Satisfaction

User satisfaction is divided into the users perceived sense of usability and usefulness of the prototype. Usability refers to if the

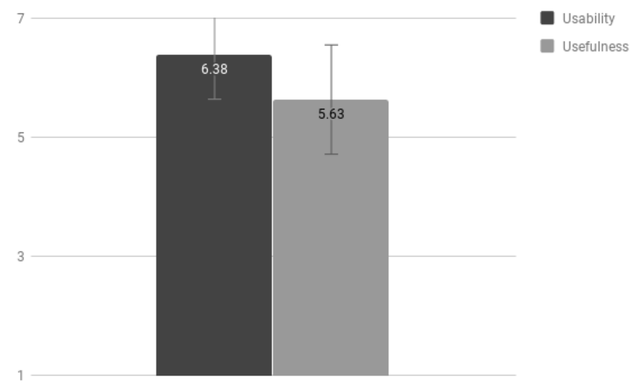


Figure 5: Mean usefulness and usability score together with standard deviation. Y-axis represents the 1 - 7 score possibilities of the UMUX-Lite survey.

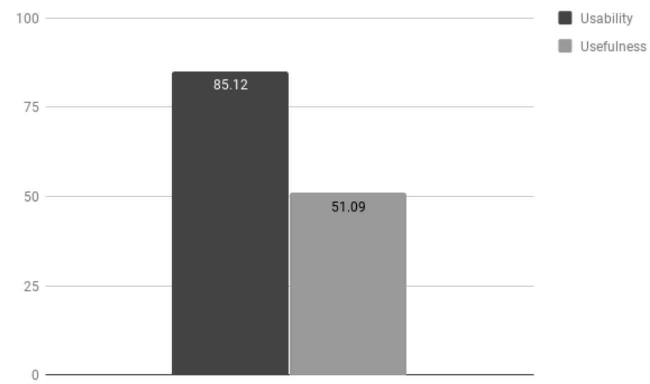


Figure 7: Percentile ranks side by side.

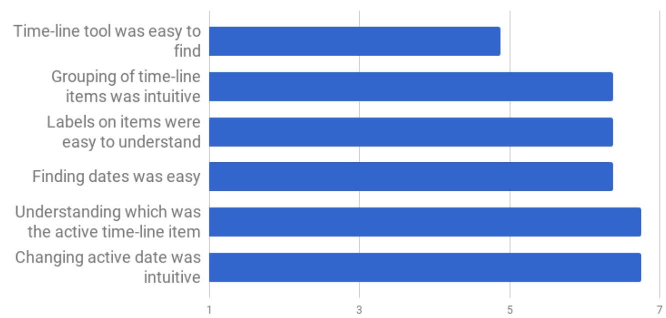


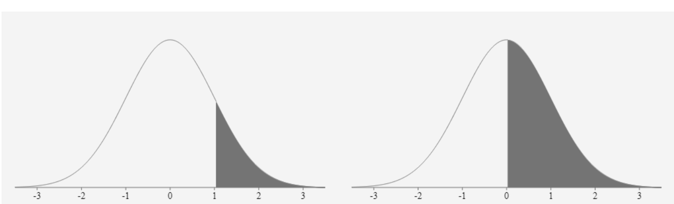
Figure 8: Additional questions asked inspired by the UMUX-Lite formulation. Shown as a mean score with standard deviation.

prototype is working. Usefulness refers to if it provides added value to the users. Usability and usefulness were measured using the UMUX-Lite survey [19]. The survey contained the standard UMUX-Lite questions but with some additional questions inquiring about specific design element (see figure 8). This was done to try and narrow down the underlying causes for the score, both positive and negative. Only the score on it own does not provide much information on where the design failed. The average usefulness and usability score can be viewed in figure 5.

The percentile rank is calculated by using a normal distribution with the mean 5.6 which represents 80% of max score of 7 [23]. The Z-score is then calculated by.

$$z = \sum_{i=0}^{n} \frac{x - \overline{x}}{\sigma}$$

The z-score can then be converted to the percentile rank using either a normal distribution table or calculator.



(a) Usability percentile rank 85%. **(b) Usefulness percentile rank 51%**

Figure 6: Percentile ranks compared to a mean score of 5.6 [23]. X-axis states the number of standard deviations away from the mean.

The usability score performed well with a percentile rank of 85%. The usefulness score however is only slightly above the mean value of the distribution with a percentile rank of 51% (see figure 6).

Answers to the additional questions that were asked in the survey (see figure 8), they were designed to shed more light on which parts of the interaction faced difficulty and affected the overall score. Results show that question 1, finding the timeline tool is significantly difficult. Especially compared to the other tasks. Both changing the date and understanding which date is selected was very well received by the users (questions 5 and 6). The midrange tasks show adequate results but with minor room for improvement (questions 2, 3 and 4).

5.2 Observations

While a user carried out tasks within the prototype observations were noted and compiled. See table 2 for observations. The highest occurring observation, 6 times out of 8, highlights the need for the timeline to convey that it can be zoomed. Further observations highlight the need for proper tooltips to be implemented, 4 found this confusing. Tooltips can seem trivial but is significantly useful for new users to help them navigate the interface. Selecting multiple time points is also an interesting observation, 4 mentioned or tried selecting multiple dates. It is not trivial how the interface should respond on selecting multiple image time points in a maps interface, see discussion section for further analysis. Having the timeline initially disabled seems to be one of the major obstacles for users. 4 people tried the left menu first, 3 tried the search box, 1 user expressed that the icon does not associate with a timeline. These provide evidence to the lack of visibility of the timeline and also the low score linked to finding the timeline. Since the timeline icon had a gray overlay while disabled it lead users to focus more on other parts of the interface initially. Causing difficulty in finding the tool. It is however important that the user does not accidentally load the whole dataset since this would cause large loading times. Which could happen if the default selected area is the whole globe. A number of observations were left out mainly due to being requests for additional features that falls outside the scope of this paper. For example 1 person claimed that the home button of the interface was unclear. However this paper is not intended as a general usability study of the GEOSS portal interface.

Table 2: Observations ordered by their frequency in descending order. Observation are only counted once per unique participant.

Frequency	Observation
75%	Ability to zoom in the timeline is not understod by the user
50%	Left side menu grabs initial attention
50%	Selecting a new geographical area should trigger loading of that area
50%	Missing tooltip causes confusion
50%	User wants to select multiple time points
50%	Understanding the purpose of the selection tool and how it is used is NOT intuitive
37.5%	User tried to use the search box when searching for the timeline tool
37.5%	The timeline icon is difficult to see on bright backgrounds
25%	Select area by dragging is requested or tried by user
25%	Legend is needed for understanding data
25%	User wants to be able to search for dates in the timeline
12.5%	User wondered if data was loading
12.5%	User expressed that the selection tool is intuitive to use
12.5%	Zooming anywhere on the timeline bar should be possible
12.5%	Timeline icon does associate with timeline
12.5%	Timeline icon does NOT associate with timeline
12.5%	The timeline disabled icon conveys that the tool should not be pressed and not that it can be enabled
12.5%	Better convey why timeline is disabled
12.5%	Having the toolbox on the right makes sense
12.5%	There should be a slider for picking date

5.3 Interview

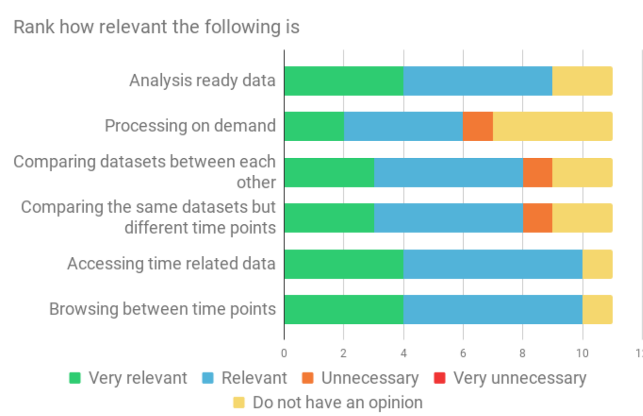
Five users appreciated the quick and responsive feel of the timeline tool. Panning and zooming caused no noticeable delays. Loading a new date only caused a brief delay that seems to be considered acceptable. 1 person voiced a negative opinion on the load time of the data, even though the selected area was fairly small. The load time was however dependent on server response time, which would fluctuate between sessions. Giving different participants a slightly different experiences.

Browsing between dates was considered sufficiently easy by 2 persons, but having the ability to search for dates was voiced 3 times as an added benefit. There were no negative opinions on browsing, 1 person claimed that navigating the timeline was easy after realizing that you have the ability to zoom. Having a slider tool to move between dates was also suggested by 1 person. The minimalistic design and the ability to hide the timeline was seen as a benefit by 1 person. One user claimed that it was good to have a lot of space for the map. 2 persons struggled with zooming in the timeline due to only being able to scroll in the top half of the timeline. This came up as an improvement 2 times.

Users were otherwise satisfied with the zooming of the timeline, 4 out of 8. Providing hints to the user that you are able to zoom in the timeline is something that came up multiple times, both in the observations and as improvement suggestions in the interview. 2 users would like to have plus and minus symbols, 1 person thought that zooming was not intuitive. Providing a legend together with the information was a point that came up 2 times. A tool to provide a data legend is already in the workings of other developers and will therefore not be considered in this report.

The timeline tool icon was on 1 occasion deemed not intuitive and on 1 occasion deemed to be intuitive. The icon should be revisited to understand how it can be improved. The ability to pan left and right in the timeline could be explained with more hints. One person suggested an improvement with left and right arrows at the far ends of the timeline. However panning left and right was deemed to work well by 3 users. Selecting multiple time points was of interest to 3 users. Giving the ability to pick multiple points raises the question on how they should be handled, what way of visualizing would be appropriate? 2 suggestions were to show the time points side by side. Comparing between picked time points was also considered to be of interest for 2 users.

Figure 9: Results gathered from from the survey sent out after the GEOSS Data Provider Workshop.



5.4 Workshop Survey

The survey inquired about how users would like to access data. For this report information relating to data cubes are included. In figure 9 it is clearly shown that a majority of the participants consider all mentioned data cube capabilities to be either relevant or very relevant. The lowest scoring capability was processing on demand, 6 out of 11 believed it to be relevant or above while 1 person found it unnecessary and 4 people had no opinion. In the survey some questions were designed to inquire about how relevant time analysis and comparison was to the users. 10 out of 11 found both accessing time related data and browsing time related data to be relevant or above. 8 out of 11 found comparing time related data to be interesting. The result servers to reinforce the idea that allowing for browsing and accessing the time dimension is highly relevant to GEOSS portal and platform users.

The answers to the open questions were quite divided. 2 out of 11 would like to compare data by processing algorithms on the server side. 2 people mentioned downloading the data and doing it themselves locally. 1 person suggested to overlay the data on the map and 1 person suggested to show the data side by side.

6 DISCUSSION

6.1 User satisfaction

The user satisfaction was reasonably high. The wide range of possible interactions in the timeline was one of the more praised attributes. Interactions include zooming, panning and selecting. This is probably a major reason for the high usability score. It was important that the user easily could get a sense of the extent of the data, where it was lacking, where it was sufficient to use. What is deemed sufficient is entirely case based and up to the scientist, the visualization purpose is only to convey precise information for the user to base a decision on. The design enabled the user in their task and interaction was not interrupted by limitations.

The usefulness score was about average. The reason seems to be the obstacle people had with finding the tool. Improving the visibility of the tool seem to be essential. Grouping the timeline items could also be improved. One suggestion is to group items based on years, months, days. Which might prove to be more intuitive. Since the grouping is done by a greedy algorithm, picking the first best options and combining them. It can seem arbitrary and often create different sized groups. Grouping in a more homogeneous way, enforcing groups of equal size could prove to be beneficial. It might be more intuitive and agreeable for the user. Finding dates could be improved according to results. A large part of this seems to be due to the timeline items not grouping directly when opening the timeline tool. Timeline items only group while zooming. Since zooming was not intuitive for many users, the result is a cluttered timeline.

6.2 Affordances

The affordances of finding the tools relies on the icon which should resemble what the users expects when looking for a timeline. The icon should also be placed in a visible position in order for the user to successfully perceive the affordance. The icon should be contrasted to the background and change color on hover, to indicate that it is a button and that it is interactive. The icon of the prototype was made to look like the items in the timeline. This could be beneficial for long-term use, when the user already knows what the timeline looks like. Creating a learned association and affordance between the icon and the tool itself. For first time users this might be less successful, the icon should then play on general associations of what timelines look like. It is an interesting balancing between catering to first time users and experienced users. Clearly the result of this study states that the icon can be improved for new users. Further development should take care not excluding experienced users in this regard.

Even though the timeline affords zooming there is a lack of visual cues to support the user in perceiving this affordance. Further development should keep in mind to create proper visual cues. These cues could be plus and minus buttons suggesting the ability to zoom. It could also be a zooming animation taking place when opening the timeline. There are a number of ways of accomplishing this. Some of the other portals that are frequently used at ESRIN do not allow for zooming in the timeline. This has probably reinforced the notion among users that zooming is not possible in timelines.

The ability to pan right and left was discovered more easily by the users. This was probably due to the users noticing that dates and timeline items went outside of the visible range. This forced the user to try and move left and right to find them. Users are familiar with vertical scrollbars on websites. Scrolling or panning when information is outside of a persons visual reach is therefore natural to users due to the learned affordance. Putting items right on the border of visibility therefore could act as a hint at additional space beyond which can be explored.

Using a gray overlay on an icon indicates that it is disabled. The user perceives that the button does not afford clicking. It seems users also tend to investigate disabled tools the very last. Prioritizing tools that they perceive to be active and usable. A solution to the problem could be to not disable the tool. Instead when the tool is opened it contains instruction of what is required for usage. This would eliminate the affordance that the tool can not be used. While also providing more explicit feedback to the user of how it is intended to be used.

6.3 Limitations

A timeline comes with the affordance that it shows time. Only showing time is a direct limitation when working with data cubes. Data cubes could potentially store any dimension of data. The horizontal bar could provide a space for other dimensions such as elevation. The date labels with months, days, etc, would have to be replaced with altitude and the grid should be reformatted from daily increments to a more fitting spacing for altitude, such as one gridline per 100 meters. The timeline seem to also afford a selection box, when users try to select multiple dates they do so by clicking and dragging. The user quickly finds that it is not possible but it can be a viable option if multi select would be implemented. One user tried to deselect a time point by clicking the point again. It is a limitation that the timeline can not afford deselecting areas.

Currently the format of received data is a large sets of pixels. A more beneficial way of loading data is to provide them by tiles. The connection can then load tiles asynchronously and start showing the regions that are loaded straight away. Tiled sources usual supply different scaled versions based on the current zoom level. When you are looking from afar the loaded image is a scaled down version. When you zoom in the full sized version is supplied.

Another limitation is that the timeline currently only supports viewing one dataset. Potentially the user would like to view multiple datasets in parallel on the timeline interface. This can be useful to understand how the selected time points relate to each other on the time axis. It can be beneficial if the user has multiple datasets that they want to interact with at the same time. It would be inconvenient to the users if they have to switch datasets frequently. Adding more datasets however requires more space. It would be a challenge to show plenty of datasets while maintaining good readability and room for interactions with each dataset.

6.4 Time range or discrete time points

While conducting the study a number of users voiced a need for selecting multiple time points. For satellite data it might be more intuitive to select data as a range instead. This relates well with the concept of selecting an area of interest that is used in the GEOSS portal for geographical area selection. Selecting a range could prove to be more intuitive than selecting plenty of discrete points. The data cube used in this study supports on the fly mosaicking of the sets of images. Which makes the task significantly easier and relieves the portal of any processing requirements. Comparing different time ranges might prove to be less intuitive but should be investigated. The metadata associated with geographical coverages is often described as a time range. Making it more practical compared to retrieving the full set of available days. If the data is presented as a continuous range it would not require the clustering algorithm either. Making the implementation more convenient and less resource intensive. Though it should be noted that the clustering algorithms did not have any significant impact on responsiveness. Using time ranges instead of discrete points has a lot of advantages. Inspiration for how to design this can be gathered from timelines coming from audio editing softwares. Where extracting parts of a sound sample is a commonly used range selection operation.

6.5 Multiple datasets in a timeline

The GEOSS portal already support loading multiple datasets and storing them in so called layers. It is easy to imagine that this layer analogy could be extended to the timeline interface. Since the timeline is a horizontal axis it could easily be stacked with multiple layers in the vertical direction. Having them stacked vertically is also the representation commonly seen in softwares that use image layers, such as photoshop. This could then be useful for the user when comparing between datasets. Because they could also get an understanding how the features they see on screen relate to each other in the time domain. Understanding if the datasets compared are one day apart or months apart can be essential for a data scientist.

6.6 Multiple dimensions

Since data cubes support in theory any number of dimensions it should be considered how this can be implemented in an interface. The problem with having a general tool that works for any dimension is that it can not have affordances related to time, since that you for example could cause confusion if a dataset containing elevation would be loaded. The affordances should switch dynamically depending on the data contained. A simple example of this is if elevation data would be loaded into the timeline. The labels containing time hints, such as year, month day would have to be switched according to the data presented. It could be switched to show kilometers, meters, etc instead. Another way of supporting multiple dimensions is to have separate interfaces depending on the data type. Elevation information could for example be shown as 3D depth information on top of the map. Another solution would be to have the vertical axis of the timeline show elevation data. Which would in practice turn the timeline into a two dimensional selection tool, in a similar fashion as the bounding box that selects a geographical area. This could be beneficial for the user to understand the relation between time and elevation.

6.7 Interactively Comparing images

Comparing images is useful for users to understand how dataset relate to each other. It can also be used for the same dataset but to understand how an area has changed over time. One method mentioned in the workshop survey was to pose images next to each other. Doing that could prove to be not so intuitive since it is overlaid on top of a map. Geographical image usually are mapped to the region they convey. If you have for example two images of Italy next to each other it implies they both can not be placed in the proper position on the map. This might however be solved by temporarily disabling the map or clearly showing that the side comparison is not related to the geographical position.

One solution could be to split the map window into two or more parts. Showing multiple maps side by side. This could work but would quickly limit the individual space for each map and therefore also limiting the users ability to interact with each map. It has also been shown previously that maps should cover 70% of the interface or above for the best usability [27].

Another way could be to show them on top of each other then using an interactive slider to reveal or hide the underlying layer. The problem with this is that you are obscuring information that the user might miss. The benefit is that the quick visual change that occurs when dragging the slider could serve to highlight changes better. Human perception has been shown to pick up movement and rapid changes better than static colors [14]. Comparing images could also be done by mathematical operations and algorithms. The EO data cube used for this study supports user defined algorithms through a WCPS protocol. This would give a lot freedom to the user but can also be rather complex to understand and take benefit from. An interface for mediating this capabilities in a simple an intuitive way is a challenge and study in its own right.

7 CONCLUSION

The current state of the timeline has proven to be a useful tool for inspecting time related data served by a data cube. The user was enabled and not hindered by the design and the available interaction tools. Multiple different interaction techniques were available for the timeline tool, including zooming, panning and selecting. The visibility of the tool was not very intuitive and requires additional attention. Zooming with the tool requires more visual cues conveying that the timeline affords zooming functionality. There are still multiple challenges and considerations in creating an interface that covers more of the capabilities that data cubes could be used for. These include multiple datasets shown in the timeline, selection of data across multiple dimensional axises and comparing time sections. It is found that the timeline is suitable for extending in the vertical axis with additional datasets or another data dimension. Arguably it is not convenient to do both in the same tool. Multiple dimensions should be handled in separate tools if you want to tailor the user interface to their unique properties.

This research can serve as a foundation for others in guiding them about what parameters are important in creating a timeline interface. Since there is a push for more data cube implementations and many stories of successful insights facilitated by time analysis of data cubes, there most probably will be more web interfaces utilizing data cubes in the future. The need for providing users with proper tools for navigating the data cube will be essential. The research can also serve as a base for further investigation into how to interact and visualize the n-dimensions that data cubes could in theory contain or more generally as part of any research that deals with interaction and visualization of a temporal dimension.

8 FUTURE WORK

In the interest to further increase the compatibility between data cubes and GEOSS, one would need to evaluate which additional aspects besides time that could be useful in the GEOSS platform. One such could be elevation, arguably there might be other important parameters to find. Another useful task would be how to interact with multiple datasets while browsing the time axis and how to interact with multiple data dimension. Visualizing the relation between these dimension could be beneficial to the user, but it is also a very difficult task. Especially to find a general way that works for all parameters.

A large area that has not been approached in this report is to allow users the processing tools of data cubes. Investigating how this could be presented in a user friendly manner. For example it could be provided through a visual programming interface, with code blocks that can be connected.

Lastly the area of how to discover the data cubes through the GEOSS portal search widget has not been dealt with. It includes the task of how the GEO DAB can convey different dimensions through the established broker system.

ACKNOWLEDGEMENT

I want to thank Joost, Andrea and Guido at ESA, both for giving me the opportunity to work on this exciting project but also for the support and guidance given throughout the whole endeavor. I also want the extend my thanks and gratitude to my academic mentor Mario and my fellow peers for giving this report the final touch it absolutely needed.

REFERENCES

[1] Joakim Adrup, Joost Van Bemmelen, Andrea Della Vecchia, and Guido Colangeli. 2018. Pioneering Steps towards Use of Data-cubes in the Global Earth Observation System of Systems. In *PV2018 Conference - Adding value and preserving data*, Esther Conway and Kate Winfield (Eds.). Science and Technology (STFC) Library and Information Service, 48–50.

[2] Katherine Anderson, Barbara Ryan, William Sonntag, Argyro Kavvada, and Lawrence Friedl. 2017. Earth observation in service of the 2030 Agenda for Sustainable Development. http://www.tandfonline.com/doi/full/10.1080/10095020.2017.1333230. *Geo Spat. Inf. Sci.* 20, 2 (2017), 77–96.

[3] Natalia Andrienko, Gennady Andrienko, and Peter Gatalsky. 2003. Exploratory spatio-temporal visualization: an analytical review. http://geoanalytics.net/and/papers/jvlc03.pdf. *Journal of Visual Languages & Computing* 14, 6 (2003), 503–541.

[4] Christian Ariza-Porras, Germán Bravo, Mario Villamizar, Andrés Moreno, Harold Castro, Gustavo Galindo, Edersson Cabera, Saralux Valbuena, and Pilar Lozano. 2017. CDCol: a geoscience data cube that meets Colombian needs. https://link.springer.com/chapter/10.1007/978-3-319-66562-7_7. In *Colombian Conference on Computing*. Springer, 87–99.

[5] Nick Babich. 2016. Mobile UX Design: Sliders - UX Planet. https://uxplanet.org/mobile-ux-design-sliders-761ce4bb2a86. (April 2016). (Accessed on 02/16/2018).

[6] Benjamin Bach, Pierre Dragicevic, Daniel Archambault, Christophe Hurter, and Sheelagh Carpendale. 2014. A Review of Temporal Data Visualizations Based on Space-Time Cube Operations. https://hal.inria.fr/file/index/docid/1006140/filename/spacetime-authorversion.pdf. In *A Review of Temporal Data Visualizations Based on Space-Time Cube Operations.* 19.

[7] P Baumann. 2017. The Datacube Manifesto. http://www.earthserver.eu/sites/default/files/upload_by_users/The-Datacube-Manifesto.pdf. (2017).

[8] Marion L Butterfield, Jay S Pearlman, and Stephen C Vickroy. 2008. A system-of-systems engineering GEOSS: Architectural approach. http://ieeexplore.ieee.org/abstract/document/4626034/. *IEEE Systems Journal* 2, 3 (2008), 321–332.

[9] Max Craglia, Jiri Hradec, Stefano Nativi, and Mattia Santoro. 2017. Exploring the depths of the global earth observation system of systems. http://www.tandfonline.com/doi/abs/10.1080/20964471.2017.1401284. *Big Earth Data* 1, 1-2 (2017), 21–46.

[10] M Femke de Jong, Marilena Oltmanns, Johannes Karstensen, and Laura de Steur. 2018. Deep convection in the Irminger Sea observed with a dense mooring array. *Oceanography* 31, 1 (2018), 50–59.

[11] EORC. 2018. Global PALSAR-2/PALSAR/JERS-1 Mosaic and Forest/Non-Forest map. http://www.eorc.jaxa.jp/ALOS/en/palsar_fnf/fnf_index.htm. (May 2018). (Accessed on 05/23/2018).

[12] Amir Gandomi and Murtaza Haider. 2015. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management* 35, 2 (2015), 137–144.

[13] Gregory Giuliani, Bruno Chatenoux, Andrea De Bono, Denisa Rodila, Jean-Philippe Richard, Karin Allenbach, Hy Dao, and Pascal Peduzzi. 2017. Building an Earth Observations Data Cube: lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). http://www.tandfonline.com/doi/abs/10.1080/20964471.2017.1398903. *Big Earth Data* 1, 1-2 (2017), 100–117.

[14] Gunnar Johansson. 1975. Visual motion perception. *Scientific American* 232, 6 (1975), 76–89.

[15] Page Laubheimer. 2017. Sliders, Knobs, and Matrices: Balancing Exploration and Precision. https://www.nngroup.com/articles/sliders-knobs/. (May 2017). (Accessed on 02/16/2018).

[16] Laurent Lefort, Josh Bobruk, Armin Haller, Kerry Taylor, and Andrew Woolf. 2012. A linked sensor data cube for a 100 year homogenised daily temperature dataset. https://www.researchgate.net/profile/Laurent_Lefort/publication/267263194_A_Linked_Sensor_Data_Cube_for_a_100_Year_Homogenised_daily_temperature_dataset/links/54f261ef0cf24eb879482f85.pdf. In *Proceedings of the 5th International Conference on Semantic Sensor Networks-Volume 904*. CEUR-WS. org, 1–16.

[17] Adam Lewis, Simon Oliver, Leo Lymburner, Ben Evans, Lesley Wyborn, Norman Mueller, Gregory Raevksi, Jeremy Hooke, Rob Woodcock, Joshua Sixsmith, Wenjun Wu, Peter Tan, Fuqin Li, Brian Killough, Stuart Minchin, Dale Roberts, Damien Ayers, Biswajit Bala, John Dwyer, Arnold Dekker, Trevor Dhu, Andrew Hicks, Alex Ip, Matt Purss, Clare Richards, Stephen Sagar, Claire Trenham, Peter Wang, and Lan-Wei Wang. 2017. The Australian Geoscience Data Cube — Foundations and lessons learned. https://www.sciencedirect.com/science/article/pii/S0034425717301086. *Remote Sens. Environ.* 202 (2017), 276–292.

[18] James R Lewis. 1994. Sample sizes for usability studies: Additional considerations. *Human factors* 36, 2 (1994), 368–378.

[19] James R Lewis, Brian S Utesch, and Deborah E Maher. 2013. UMUX-LITE: when there's no time for the SUS. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2099–2102.

[20] Joanna McGrenere and Wayne Ho. 2000. Affordances: Clarifying and Evolving a Concept. http://teaching.polishedsolid.com/spring2006/iti/read/affordances.pdf.

[21] Stefano Nativi, Massimo Craglia, and Jay Pearlman. 2012. The Brokering Approach for Multidisciplinary Interoperability: A Position Paper*. https://www.researchgate.net/profile/Stefano_Nativi/publication/258625711_Integrated_brokering_framework_for_multi-disciplinary_research/links/57a363d708ae3f4529220765/Integrated-brokering-framework-for-multi-disciplinary-research.pdf. *International Journal of Spatial Data Infrastructures Research* 7 (2012), 1–15.

[22] Stefano Nativi, Paolo Mazzetti, Mattia Santoro, Fabrizio Papeschi, Max Craglia, and Osamu Ochiai. 2015. Big data challenges in building the global earth observation system of systems. https://www.sciencedirect.com/science/article/pii/S1364815215000481. *Environmental Modelling & Software* 68 (2015), 1–26.
[23] Jakob Nielsen and Jonathan Levy. 1994. Measuring usability: preference vs. performance. *Commun. ACM* 37, 4 (1994), 66–75.
[24] Donald A Norman. 1999. Affordance, conventions, and design. https://dl.acm.org/citation.cfm?id=301168. *interactions* 6, 3 (1999), 38–43.
[25] Group on Earth Observation. 2018. GEOSS portal. http://www.geoportal.org//. (April 2018). Accessed: 2018-04-18.
[26] Bernd Resch, Florian Hillen, Andreas Reimer, and Wolfgang Spitzer. 2013. Towards 4D Cartography – Four-dimensional Dynamic Maps for Understanding Spatio-temporal Correlations in Lightning Events. http://www.tandfonline.com/doi/abs/10.1179/1743277413Y.0000000062. *Cartogr. J.* 50, 3 (2013), 266–275.
[27] Bernd Resch and Bastian Zimmer. 2013. User Experience Design in Professional Map-Based Geo-Portals. http://www.mdpi.com/2220-9964/2/4/1015/htm. *ISPRS International Journal of Geo-Information* 2, 4 (2013), 1015–1037.
[28] Peter Strobl, Peter Baumann, Adam Lewis, Zoltan Szantoi, Brian Killough, Matthew Purss, Max Craglia, Stefano Nativi, Alex Held, and Trevor Dhu. 2017. The Six Faces of the Data Cube. https://www.researchgate.net/profile/Zoltan_Szantoi/publication/321385814_The_Six_Faces_of_the_Data_Cube/links/5a200deaaca272088b23f822/The-Six-Faces-of-the-Data-Cube.pdf. In *The Six Faces of the Data Cube*. 4.
[29] USGS. 2016. Land Change Monitoring, Assessment, and Projection. http://www.earthobservations.org/documents/meetings/201605_lcc/201605_lcc_usgs_climate_land_use_change.pdf. (January 2016). (Accessed on 02/19/2018).
[30] Joost Van Bemmelen, Guido Colangeli, Stefano Nativi, Mattia Santoro, Roberto Roncella, Gregory Giuliani, Elrich Frazier, Michelle Antony, Paola De Salvo, and Wenbo Chu. 2017. The GEOSS Platform Manual. https://www.earthobservations.org/documents/gci/201711_gci_manual_01.pdf. (November 2017). (Accessed on 05/14/2018).
[31] Curtis E Woodcock, Richard Allen, Martha Anderson, Alan Belward, Robert Bindschadler, Warren Cohen, Feng Gao, Samuel N Goward, Dennis Helder, Eileen Helmer, et al. 2008. Free access to Landsat imagery. http://science.sciencemag.org/content/320/5879/1011.1. *Science* 320, 5879 (2008), 1011–1011.
[32] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. https://dl.acm.org/citation.cfm?id=1240704. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*.

APPENDIX

Interview Answers

Interview answers shortened and summarized to retrieve their essence. If the same answer came up on multiple occasions it is only noted once but with a number in front showing the frequency of that answer.

(1) What parts worked well?
- 5 Responsiveness of the timeline and loading of data
- 5 Timeline in general
- 4 Zooming in timeline
- 3 Scrolling in timeline
- 2 Easy to move between dates
- 1 Scaling of time ranges, i. e. year, month, day
- 1 Selecting area
- 1 Finding area
- 1 Design, aesthetics vice
- 1 Leaving a lot of space for the map

(2) What parts worked poorly?
- 4 More tooltips
- 2 Legend should be available
- 2 Plus and minus button to show that you can zoom in the timeline
- 2 Zooming should be enabled on the whole timeline bar
- 1 Should be able to search for dates in the search box and have them show up in the timeline
- 1 Ambiguity between searching for dates in the search box and in the timeline
- 1 Slow load time of map data
- 1 Could visualize better which dates have missing data
- 1 Selecting area
- 1 Being able to zoom was not intuitive
- 1 Timeline items should be grouped directly when opening the tool
- 1 Better convey why timeline tool is disabled and how to enable
- 1 Search box should focus on search area
- 1 Visibility of the timeline icon over a bright background

(3) Did something work in another way than you initially though?
- 2 Deselecting an area should remove the data points in that area
- 1 Opening the timeline tool
- 1 Timeline icon did not convey what it was
- 1 Panning should have arrow to indicate that it is possible
- 1 Selecting worked by left clicking twice, not dragging a box

(4) What additional things would you like to be able to do with the tools?
- 3 Multiple time points overlaid
- 2 Download data subset
- 2 Show open time point next to each others
- 2 Compare variables between time points
- 1 Search for date
- 1 Show difference between time points
- 1 Save screen shot for presentations
- 1 Information about the selected data, shown on mouse hover over data
- 1 Give the user the option to chose image interpolation, allow for point interpolation instead of standard bilinear
- 1 Slider for picking the date
- 1 Overview of what data is available

User Scenarios

Who is the user?	Why do they visit geoportal?	What are their goals?	Why data cubes?	Resource
African Association of Remote Sensing of the Environment (AARSE)	Large focused on agriculture. They have investigated the use of NDVI with sentinel 2 imagery.	Inform African governments about EO systems and geo-information.	DC could provide agriculture information over time. As shown with the Australian DC.	http://www.africanremotesensing.org/
AfriGEOSS	Agriculture and food, water resources, land cover.	EO for policy and sustainable societal impact in Africa	Predicting agriculture growth or water access by looking at historical data could be interesting.	http://www.earthobservations.org/me_201706_afrigeoss.php
AtlantOS	Baryometry, salinity, sea level monitoring, etc.	The vision of AtlantOS is to improve and innovate Atlantic observing by using the Framework of Ocean Observing to obtain an international, more sustainable, more efficient, more integrated, and fit-for-purpose system.	Time	https://www.atlantos-h2020.eu/project-information/
Centre for Development and Environment (CDE)	Gather research info on deforestation Madagascar, land degredation and regeneration Mongolia, future research projects.	Geographic information systems (GIS) and satellite remote sensing are integral parts of CDE's research and development approach	Easy access to time related changes. Inherently important to investigate degredation, regeneration and deforestation.	http://www.cde.unibe.ch/services/geoinformatics/geoinformatics_research/index_eng.html
CNR Earth and Environment	Bio-diversity, Snow and ice coverage	The Department of Earth systems science and environmental technologies co	Time	https://www.cnr.it/en/thematic-areas/energy-environment
Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI)	Hydrologic systems – precipitation, streamflow, groundwater levels, water lost through evaporation, etc.	Develop infrastructure and services for the advancement of water science in the United States.	Temporal dimension for research that requires it. Could also provide a source for example where precipitation and streamflow is combined.	http://data.cuahsi.org
European Marine Board	Quote "Improve knowledge of the 4-Dimensional Ocean (a volume that changes in the time dimension) and its role in the earth and climate system, including the human component"	Seas and ocean science	Time	http://www.marineboard.eu/
Future Earth	Water, energy, food availability. Safe-guard natural assets. Decarbonise atmosphere. Among a lot of other goals.	Research for global sustainability	Changes over time to check how their goals are progressing.	http://www.futureearth.org/
GEO BON	Spread of species, bio-diversity, monitoring.	Improve the acquisition, coordination and delivery of biodiversity observations and related services to users including decision makers and the scientific community.	Track how spread progresses	http://geobon.org/
GEO Cold Regions (GEOCRI)	Data about snow cover, snow change, climate change.	Coordinated EO to facilitate well-informed decisions and support the sustainable development of the cold regions globally.	Changes require investigating the time aspect.	https://www.earthobservations.org/activity.php?id=114
Geohazard Supersites and Natural Laboratory initiative (GSNL)	Optical and/or SAR and ground-based geophysical data sets derived from different sources and different disciplines. (e.g., Seismic, GNSS, Strain meter, Tilt, Gas, gravity, LIDAR).	Aiming to improve, through an Open Science approach, geophysical scientific	Research areas of interest before, during and after geological hazards and events.	http://geo-gsnl.org/supersites/permanent-supersites/
Global Forest Observation Initiative (GFOI)	Optical and Radar data, NDVI, forest change	Provide sustained availability of forest monitoring systems	To see changes in forests, you need to compare different time points.	http://www.gfoi.org/
Global Observation System for Mercury (GOS4M)	Mercury	GOS4M will provide mercury speciation measurements (Gaseous Elemental M	Time	http://www.gos4m.org/scope/
International Ocean Carbon Coordination Project (IOCCP)	Surface CO2 levels, ocean interior observations, ocean acidification	The IOCCP promotes the development of a global network of ocean carbon observations	Time and a third spatial axis for ocean interior research.	http://www.ioccp.org/
The System of Environmental Economic Accounting (SEEA)	Gather information on thematic accounts, agriculture, fisheries, forestry, air emissions, energy, environmental activity, ecosystem, land, material flow, water	SEEA ecosystem accounting is a framework to account for ecosystem assets and ecosystem services.	DC could provide all of their accounts into one centeralized data sources, where different layers provide different accounts.	https://seea.un.org/
Belmont Forum	Bio-diversity, disaster risk, climate prediction, etc	The Belmont Forum is an international partnership that mobilizes funding of er	Time, view changes	http://www.belmontforum.org/impact/
Bioversity International	Deforestation, bio-diversity, agriculture	Research for development in agricultural and tree biodiversity	Time, view changes	https://www.bioversityinternational.org/
International Institute for Sustainable Development (IISD)	Water, energy, resilience, economic law and policy	IISD's programs conduct rigorous research and engage citizens, businesses a	Time, view changes	https://www.iisd.org/programs
The International Centre for Integrated Mountain Development (ICIMOD)	Cryosphere, climate change, hazards disaster, natural resource managment, mountain livelihoods	We support regional transboundary programmes through partnership with regi	Time, view changes	http://www.icimod.org/?q=info

Survey Results

Hackaton survey									
Age (years)	What role did you	Analysis ready da	Processing on de	Comparing datas	Comparing the sa	Accessing time r	Browsing betwee	What would be y	What would be your ideal way to compare satellite data?
25 - 34	data scientist	Very relevant	Relevant	Do not have an o	Do not have an c	Relevant	Relevant	to not download	depends on the data but I like histograms and numeric values comparison
18 - 24	developer	Very relevant	Do not have an c	Relevant	Relevant	Relevant	Relevant	Through an easy	Don't know
18 - 24	Web application	Very relevant	Do not have an c	Relevant	Relevant	Relevant	Relevant	A simple search,	Maybe, advice about how to convert one data to another, and higlighting main differences would be great
25 - 34	Developer	Relevant	Very relevant	Relevant	Relevant	Relevant	Relevant	A unified API for	Having an API which can process requests on the server before. This is not the purpose of a data broker, though.
25 - 34	Developer	Relevant	Relevant	Very relevant	Very relevant	Very relevant	Very relevant	via portals, ftps o	via algorithms applied to the respective datasets
25 - 34	Developer	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	GEOSS Portal	Image processing via third party software
25 - 34	Planner & EO da	Do not have an o	Do not have an c	Very relevant	Very relevant	Very relevant	Very relevant	I'm already satisf	I'm already satisfied with current options.
18 - 24	Developer	Do not have an o	Do not have an c	Do not have an o	Do not have an c	Do not have an c	Do not have an o	-	-
25 - 34	Developer	Relevant	Unnecessary	Unnecessary	Unnecessary	Very relevant	Very relevant	A properly impler	My own scripts.
35 - 44	Frontend develop	Relevant	Relevant	Relevant	Relevant	Relevant	Relevant	I dont have sooo	Not sure what you mean with compare? You can use just WMS for displaying the different datasets next to each other.
25 - 34	/	Very relevant	Very relevant	Very relevant	Very relevant	Very relevant	Very relevant	As easy to use a	Overlay
User study survey									
Ids	Time-line tool wa	The time-line is e	The time-line cap	Grouping of time-	Labels on items v	Finding dates wa	Understanding w	Changing active	Gender
1	5	7	5	6	5	7	7	7	Male
2	2	7	5	5	6	7	7	7	Male
3	6	6	6	6	6	7	6	7	Male
4	5	6	6	7	7	5	7	7	Male
5	6	7	6	7	7	6	7	6	Male
6	6	6	6	7	7	6	7	7	Male
7	5	7	4	7	7	7	7	7	Female
8	4	5	7	6	6	6	6	6	Female
Observations									
Person	1	2	3	4	5	6	7	8	total
Zooming of timeli	0	1	1	1	1	1	0	1	6
Left side menu gi	1	1	1	0	0	1	0	0	4
New area selecti	1	1	1	0	1	0	0	0	4
Tooltip missing c	1	1	0	1	1	0	0	0	4
Would like to sele	1	0	0	0	0	1	1	1	4
Selection box is N	0	1	1	1	1	0	0	0	4
User tried to use	1	1	0	0	0	1	0	0	3
Timeline icon is c	0	0	1	1	0	0	0	1	3
Select area by dr	0	0	0	0	1	0	1	0	2
Legend for under	0	0	0	1	1	0	0	0	2
Search for dates	0	0	0	1	1	0	0	0	2
Wondered if data	1	0	0	0	0	0	0	0	1
Selection box is i	0	0	0	0	0	0	0	1	1
Zooming on whol	0	0	0	1	0	0	0	0	1
Icon does associ	0	0	0	0	0	0	1	0	1
Icon does NOT a	0	1	0	0	0	0	0	0	1
Change disabled	0	0	0	0	0	0	1	0	1
Convey why time	0	0	0	0	0	0	1	0	1
Home button unc	0	1	0	0	0	0	0	0	1
Toolbox on right	0	0	1	0	0	0	0	0	1
Bigger boxes to c	0	0	1	0	0	0	0	0	1
Slider to pick dat	0	0	0	0	0	1	0	0	1

Study Questionnaire

Study Questionnaire

*Required

1. **Gender** *
   Mark only one oval.
   - Female
   - Male
   - Prefer not to say

2. **Age (years)** *
   Mark only one oval.
   - < 18
   - 18 - 24
   - 25 - 34
   - 35 - 44
   - 45 - 54
   - 55 - 64
   - 65 - 74
   - > 74

3. **Time-line tool was easy to find** *
   Mark only one oval.

	1	2	3	4	5	6	7	
Disagree fully								Agree fully

4. **Grouping of time-line items was intuitive** *
   Mark only one oval.

	1	2	3	4	5	6	7	
Disagree fully								Agree fully

5. **Labels on items were easy to understand** *
   Mark only one oval.

	1	2	3	4	5	6	7	
Disagree fully								Agree fully

6. **Finding dates was easy** *
   Mark only one oval.

	1	2	3	4	5	6	7	
Disagree fully								Agree fully

7. **Changing active date was intuitive** *
   Mark only one oval.

	1	2	3	4	5	6	7	
Disagree fully								Agree fully

8. **Understanding which was the active time-line item was intuitive** *
   Mark only one oval.

	1	2	3	4	5	6	7	
Disagree fully								Agree fully

9. **The time-line capabilities meet my requirements.** *
   Mark only one oval.

	1	2	3	4	5	6	7	
Disagree fully								Agree fully

10. **The time-line is easy to use** *
    Mark only one oval.

	1	2	3	4	5	6	7	
Disagree fully								Agree fully

GEOSSHACK Questionnaire

GEOSSHACK tools and platform feedback

*Required

1. **Age (years)** *
Mark only one oval.

- ( ) < 18
- ( ) 18 - 24
- ( ) 25 - 34
- ( ) 35 - 44
- ( ) 45 - 54
- ( ) 55 - 64
- ( ) 65 - 74
- ( ) > 75

2. **Rank your previous experience using satellite data** *
Mark only one oval.

	1	2	3	4	5	6	7	
None	( )	( )	( )	( )	( )	( )	( )	Expert

3. **What role did you take in your group?** *
Developer, planner, data scientist, etc.

____

4. **Rank how easy the following is** *
Mark only one oval per row.

	Very Difficult	Difficult	Decent	Easy	Very easy	Did not use
Finding data in the GEOSS portal is	( )	( )	( )	( )	( )	( )
Interacting with GEOSS portal interface is	( )	( )	( )	( )	( )	( )
Using the GEO DAB RESTful API is	( )	( )	( )	( )	( )	( )
Using the GEO DAB JavaScript API is	( )	( )	( )	( )	( )	( )
Accessing data that I need is	( )	( )	( )	( )	( )	( )
Implementing the GEOSS search box widget is	( )	( )	( )	( )	( )	( )

5. **Feedback on the GEOSS Portal**
We would be glad to hear about any inconveniences. Skip if you didn't use this

____
____
____
____
____

6. **Feedback on the APIs from the GEOSS Portal**
We would be glad to hear about any inconveniences. Skip if you didn't use this

____
____
____
____
____

7. **Rank how relevant the following services are** *
For potential improvements and new features
Mark only one oval per row.

	Very relevant	Relevant	Unnecessary	Very unnecessary	Do not have an opinion
Filter data and search result by an bounding box	( )	( )	( )	( )	( )
Near-real-time data	( )	( )	( )	( )	( )
Analysis ready data	( )	( )	( )	( )	( )
Long term data archives	( )	( )	( )	( )	( )
Processing on demand	( )	( )	( )	( )	( )
Preset algorithms to apply on data	( )	( )	( )	( )	( )
Notification on new satellite acquisition on area of interest	( )	( )	( )	( )	( )
Track satellites positions	( )	( )	( )	( )	( )
Combining datasets on the fly	( )	( )	( )	( )	( )
Comparing datasets between each other	( )	( )	( )	( )	( )
Comparing the same datasets but different time points	( )	( )	( )	( )	( )
Accessing time related data	( )	( )	( )	( )	( )
Browsing between time points	( )	( )	( )	( )	( )
Mobile version of the portal	( )	( )	( )	( )	( )

8. **What would be your ideal way to access satellite data?** *

9. **What would be your ideal way to compare satellite data?** *

10. **General feedback and comments :)**

Powered by
Google Forms

Questions

What parts worked well?

What parts worked poorly?

Did something work in another way than you initially though?

> What made you think it worked differently? Button, design, etc?

What additional things would like to be able to do with the tools?

> With multilayered data?

Scenarios:

Task 1:

Open timeline

Look for 21 of October

Select and load data from the given date

Task 2:

Select Rome area (with selection tool).

Select the date 2017-11-04

Find a nearby date that covers the whole interest region

Task 3:

Select 2017-01-23

Find a nearby date that covers the whole interest region

Close timeline