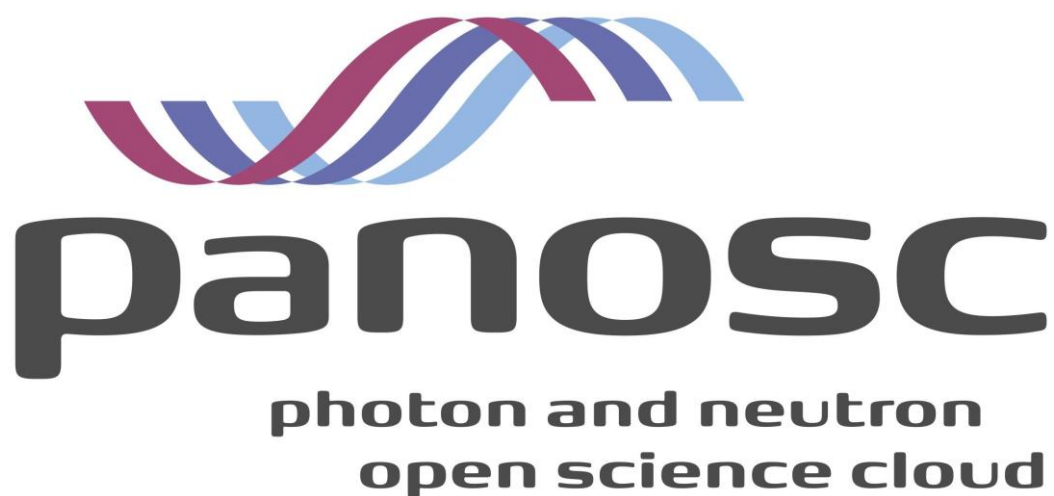


PaNOSC
Photon and Neutron Open Science Cloud
H2020-INFRAEOSC-04-2018
Grant Agreement Number: 823852



Deliverable:
D7.3 PaN EOSC Business model reference document

Project Deliverable Information Sheet

Project Reference No.	823852
Project acronym:	PaNOSC
Project full name:	Photon and Neutron Open Science Cloud
H2020 Call:	INFRAEOSC-04-2018
Project Coordinator	Andy Götz (andy.gotz@esrf.fr)
Coordinating Organization:	ESRF
Project Website:	www.panosc.eu
Deliverable No:	D7.3
Deliverable Type:	Report
Dissemination Level	Public
Contractual Delivery Date:	31/05/2022
Actual Delivery Date:	05/10/2022
EC project Officer:	Flavius Pana

Document Control Sheet

Document	Title: PaN EOSC Business model reference document
	Version: 1
	Available at: https://github.com/panosc-eu
	Files: 1
Authorship	Written by: Angela Zennaro, Giuseppe La Rocca, Ornella de Giacomo and Teodor Ivanoaica
	Contributors: Dario Roccella, Dariusz Brzosko, Diego Scardaci, Fabio Dall'Antonia, Florian Gliksohn, Jordi Bodega, Erwan Le Gall and Thomas Holm Rod
	Reviewed by: Andy Götz, Jordi Bodega
	Approved: Andy Götz

List of participants

Participant No.	Participant organisation name	Country
1	European Synchrotron Radiation Facility (ESRF)	France
2	Institut Laue-Langevin (ILL)	France
3	European XFEL (XFEL.EU)	Germany
4	The European Spallation Source (ESS)	Sweden
5	ELI European Research Infrastructure Consortium (ELI-ERIC)	Belgium
6	Central European Research Infrastructure Consortium (CERIC-ERIC)	Italy
7	EGI Foundation (EGI.eu)	The Netherlands

Table of Content

Project Deliverable Information Sheet	2
Document Control Sheet	2
List of participants	2
Table of Content	3
1. Introduction	4
2. Executive summary	5
3. Current situation	6
3.1 Data and computing in the current context of research facilities	6
3.2 PaNOSC key outcomes	6
3.3 Current vision to access infrastructure, data and data services	7
3.4 Current challenges	8
4. Methodology	10
4.1 Why considering new business models is necessary	10
4.2 Business models' considerations	11
4.2.1 The business model canvas	11
4.2.2 The multi-sided value proposition canvas	11
5 Business models	13
5.1 Business models for researchers	13
5.1.1 Value proposition	13
5.1.2 Business Model Canvas	15
5.2 Business models for commercial users	17
5.2.1 Business model I: Complete environment for data search, analysis, simulation and storage	18
5.2.1.1 Value proposition	18
5.1.2.2 Business Model Canvas	19
5.2.2 Business model II: Improving company's staff skills on data analysis and access to related tools	20
5.2.2.1 Value proposition	20
5.2.2.2 Business Model Canvas	21
5.2.3 Business model III: Complete consultancy service	22
5.2.3.1 Value proposition	22
5.2.2.3 Business Model Canvas	23

1. Introduction

Since 2019 the European Open Science Cloud (EOSC) ESFRI cluster projects like PaNOSC (<https://panosc.eu>) and ExPANDS (<https://expands.eu>) have worked on developing and supporting the adoption of the building blocks of a data commons by the Photon and Neutron (PaN) scientific community. These data commons, ranging from policies and data services structuring the support of the management to common data standards adopted, maintained and continuously improved by a scientific community which is now starting to also share services for computing and processing FAIR data from the PaN RIs.

Some of the developments towards Open Science at facilities were co-funded by the European Commission (EC) through EOSC projects (e.g. PaNOSC), while others were covered with the regular funding of facilities who had committed to provide open and FAIR data, integrating this as one of their core activities, aligned with the mission of their Research Infrastructures (RIs).

In order to sustain all the developments achieved during these years thanks to the EC funding provided to PaNOSC the RIs will require funding to maintain, operate and further Improve these developments. This document will present business and funding models that could be applicable to the RIs to sustain the outcomes of PaNOSC.

2. Executive summary

EC funded EOSC projects like PaNOSC have enabled many developments towards Open Science and FAIR data¹ at the participating RIs in order to achieve the EOSC vision of a "Web of FAIR Data and Services".

The ambition of the EOSC² is to provide researchers, Innovators, companies and citizens with a federated and open multi-disciplinary environment where they can publish, find and re-use data, tools and services for research, Innovation and education purposes. PaNOSC shares the vision of the EOSC, however, in practical terms making this vision a reality requires resources, both infrastructure (e.g. networks, computing) and personnel, that are limited and in some cases were not foreseen in the budgets of the RIs.

This document presents the key outcomes from PaNOSC that would enable the vision of the EOSC to be achieved, the associated challenges, cost structures and possible funding models together with several business models to clearly highlight all the aspects to be considered for a successful PaN EOSC business plan.

¹ Findable, Accessible, Interoperable and Reusable. <https://www.go-fair.org/fair-principles/>

² <https://eosc-portal.eu/about/eosc>

3. Current situation

3.1 Data and computing in the current context of research facilities

The project PaNOSC has allowed 5 ESFRI facilities, 1 ERIC and 1 E-Infrastructure to work together for a period of 4 years. During the project's lifetime, these infrastructures (RIs) agreed on a framework data policy to be implemented, with the objective to comply with FAIR principles, making data produced Findable, Accessible, Interoperable and Reusable. Achieving the implementation of FAIR principles implied several changes at the level of RIs, that went from the modification of their control systems to acquire relevant metadata, the adoption of standard formats for data, the use of electronic logbooks, jupyter notebooks and other tools and workflows for data analysis. This important step is supporting the reproducibility of scientific experiments and scientific results, interdisciplinary research, and, by promoting the reuse of data it improves scientific experiments, the evolution of common data culture, common training and preparation of new generations of researchers, thus maximizing the value and impact of research facilities' data by putting FAIR data in the foreground.

In this context, leveraging the data capabilities of the research facilities and promoting FAIR principles among research facilities and users, the EOSC is addressing both the reproducibility³ of the results, promote open science standards and services and, overall delivering higher quality metadata and data for better science.

3.2 PaNOSC key outcomes

The 12 main outcomes of PaNOSC delivered by the project that would be desirable to maintain in the future were grouped together into 6 Key Exploitable Results (KERs) in a survey for the EOSC Association as follows:

KER 1: FAIR-compliant Data Policy Framework for Research Infrastructures -

<https://doi.org/10.5281/zenodo.3862701>

PaNOSC FAIR data policy is an update of the PaNdata Data Policy Framework published in 2010. The PaNdata policy framework has been used as a blueprint for the open data policies of the majority of photon and neutron sources in Europe over the last 10 years. The PaNOSC Data Policy Framework will hopefully have a similar success.

KER 2: Federated Search API + data portal

<https://data.panosc.eu>

One of the main objectives of PaNOSC was to provide open data for the EOSC. The federated search API and portal aims to provide an easy way to search for open data across the photon and neutron facilities via a single portal. A recent example of such data is the Human Organ Atlas (<https://human-organ-atlas.esrf.eu>) which will provide an atlas of all human organs scanned at the micron level.

KER 3: VISA+ web viewers for HDF5 (H5Web, HDF5 Nuvola)

<https://www.panosc.eu/services/data-analysis/> and <https://github.com/silx-kit/h5web>

The VISA remote desktop platform is one of the main developments produced by PaNOSC. It was used for remote experiments during the COVID-19 confinement period and for remote data analysis since then. It provides a remote desktop experience which enables desktop sharing between multiple users and local experts. It can be deployed on a common cloud infrastructure as a general purpose solution for remote data analysis or on a remote cloud infrastructure.

KER 4: Community AAI

<https://umbrellaid.org/what.html>

³ <https://www.nature.com/articles/533452a>

The UmbrellaID community AAI is based on the AARC blueprint architecture and ready for the EOSC AAI federation. PaNOSC is working very closely with GEANT to ensure that UmbrellaID will be one of the first community AAI to form the EOSC AAI federation. The community AAI is used for community services like the metadata catalogues, training platform and software catalogue.

KER 5: Community e-learning platform

<https://e-learning.pan-training.eu>

The training and e-learning platform is an updated version of the training platform developed during the H2020 project SINE2020. PaNOSC has joined forces with the ExPaNDS project in delivering a single solution for the PaN community, the training catalogue and e-learning platform.

KER 6: Simulation of experiments (VINYL which includes SIMEX, McSTAS and OASYS)

<https://www.panosc.eu/services/data-analysis-simulation-data-system/>

McSTAS and OASYS are the de facto tools used for designing beamlines world-wide. SIMEX builds on the outcomes of the EUALL project and in addition to the previous, allows to simulate the interaction radiation-matter.

In general, the business models for the PaNOSC outcomes which are not integrated in the mission of the facilities e.g. compute resources for open data, should cover all possibilities, from services for free to full cost recovery, with all the possibilities in between. As mentioned before, the models with cost recovery are not common and for sure not popular, but on the other hand, if funding is not available or not enough to maintain a certain level of operations, new strategies need to be devised.

3.3 Current vision to access infrastructure, data and data services

Over the years, computing, simulations and scientific experiments have become equally important for scientists around the world, and many scientists, from different communities, see “The computer — your Virgil in the world of atoms⁴” as the extra capacity, connected to experiments, used to maximise the efficiency of scientific instruments. It is not unusual to have this IT and computing support ecosystem evolve as the “+1” experiment of a research facility, improving data aspects, facilitating the computation capabilities and accelerating science.

From the computing and data perspective, a research facility is a “big data” facility, a specialised (scientific and, in many cases also engineering) data producer. In this context, it is extremely important for a scientific research facility to maximise the value and the use of its most important asset, its scientific data.

The data and computing environment is the support layer facilitating experiments, accelerating data analysis processes and, most importantly, helping users better understand the experiments, the scientific instruments and scientific setups, helping them to improve the overall quality of the research performed. In this context, the facility has the full details and documentation about the data standards, about the facility experimental and computing capabilities and, in the Open Science context, can provide computing capacity as a service, as an extra experiment for users. However, while facilities are prepared to offer open and unlimited⁵ access to data, the computing capacity, as well as the support needed to make the best use of it, is limited and can only be offered to a selected group of users with the current funding schemes.

Open and FAIR Data and Computing have the potential to become the most important tool leveraged to prepare and improve users’ experiments. A Research Infrastructure, providing users access to advanced instrumentation for experiments, can consider providing access to dedicated resources, dedicated computing

⁴https://s3.eu-de.cloud-object-storage.appdomain.cloud/kva-image-pdf/assets/globalassets-priser-nobel-2013-kemi-rattigheter-press_ke_en_13.pdf

⁵ According to the current demand. The effect of EOSC in the access to data stored at the facilities is still an open question.

capacity as part of a users' campaign. This can either be included in the experimental plan as part of the proposal, or may be structured as a parallel call for proposals for computing and data capabilities, in a similar way to beam time. This allows users to apply for computing resources and data analysis support, just like normal experiments. In this context, the user would get dedicated computing capacity, resources and a team able to support the analysis and/or the specific simulations required and facilities would enhance the knowledge transfer and contribute to the creation of a common data culture across different scientific communities.

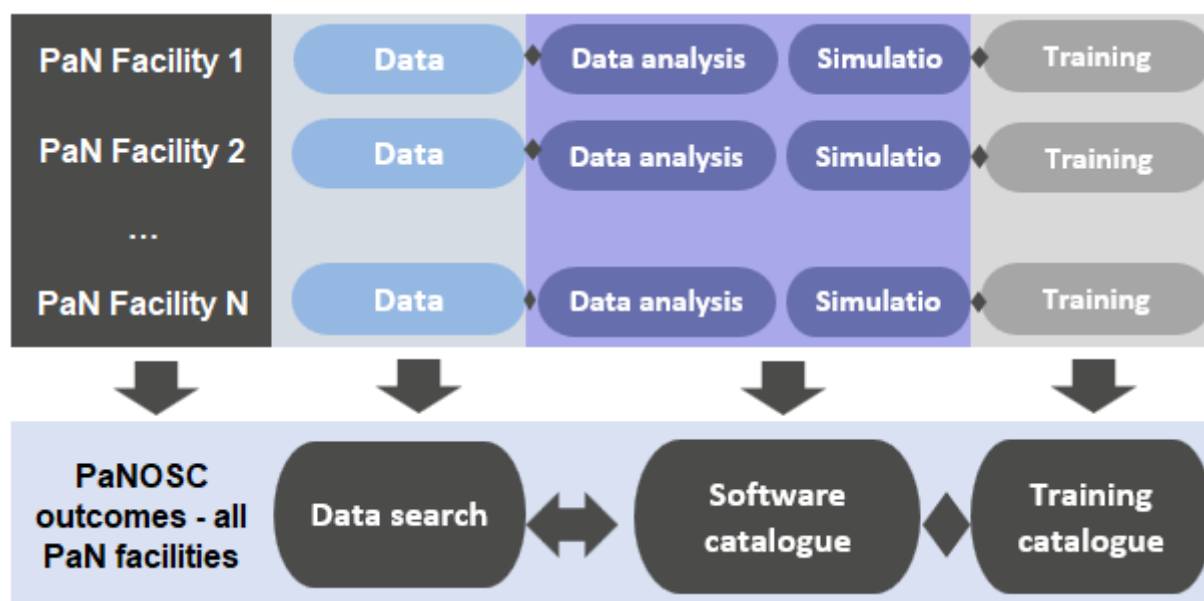


Figure 1: Representation of the current services offered by RIs and the new services that derive from FAIR data and services. While access to data may be unlimited, the capacity for other IT services such as simulation, data analysis, storage and training is limited, and therefore access to these resources should be granted according to some kind of priority allocation.

In the hypothesis of opening the call for access to tools and services for data analysis and simulation also to other researchers interested in the same resources, not limited only to users of the instruments of the facility, the dimension of the data and services infrastructure in research facilities would become as important as the access to the instruments. This process would support the transfer of know-how in both ways: facilities would train users on FAIR Data and Services and, at the same time, users would contribute with their knowledge on computing methods used for data analysis in their specific area. The development of new solutions will further enrich the pool of services of facilities and contribute to the development of EOSC, where these services would be offered as well.

3.4 Current challenges

Providing access to resources for the users of a facility (intended as those using the instrumentation such as beamlines and neutron instruments) is highly beneficial for the facility, as discussed above. However, some aspects need to be considered:

- Facilities do not have the computing capacity and resources necessary to serve all users of the EOSC as, in most cases, the resources for external users are not covered by the operations budget of the facilities. In the best case, some facilities have foreseen increasing their budget for these activities, to serve their own users, but have a very limited capacity to offer to generic EOSC users. In this case, the capacities will be very likely used as a proof of concept;

- Facilities make available standards and services but, even if these are published, users may not be familiar with computing and analysis. Training users and creating the necessary culture is crucial and it's a challenge which will be addressed over many years. Dedicated computing resources are required to train users and promote standards, resources that, again, might not have been included in the regular operational costs of the facility.
- Users, applying for experiments, or considering applying for experiments, might be interested in running simulations or analysis of existing datasets to support a possible proposal for an experiment. This is in the interest of the facility, but without dedicated computing capacities for this purpose, the optimisation of experiments based on simulations or analysis of pre-existing data is not possible.
- Training users, existing or future, requires dedicated resources, computing and, in some cases even a support team. These activities should be considered as regular operations of a research facility but, in many cases these resources might not be available.
- Some public funded RIs have hard limits on the amount of income coming from commercial activities (e.g. < 10% of income of the total budget for the ESRF) within their current legal structures, therefore new sources of income from the business models to be presented are limited.

Funding based on grants allow the development of new services and new collaborations, fundamental for Open Science and for developing common standards in a community. But grants are limited in time and scope, and the regular funding of facilities tends to be limited to a fixed amount, therefore there is the need to elaborate new business models to maintain the collaborations and the outputs updated and operational.

4. Methodology

4.1 Why considering new business models is necessary

Some of the developments towards open science at facilities were co-funded by the EC through projects (e.g. PaNOSC), while others were covered with the regular funding of facilities who had committed to provide open and FAIR data, integrating this as one of their core activities, aligned with the mission of their RIs. In other cases, the governing bodies have not yet adopted the provision of FAIR and open data and data services as part of the core mission of the facilities. Some member states have produced national policies for the implementation of the EOSC but not all of them have allocated proper funding for these activities. This is expected to change in the near future as the EOSC develops and better alignment between national and European policies is achieved. The current situation is that many RIs managers and funders don't have a full picture of the efforts involved in providing FAIR and open data. Some funding linked to EOSC activities at the national level is being assigned, but RIs play a role that is not fully understood yet, in part because the provision of open and FAIR data and the related services is something relatively new for some RIs. On one hand, it seems straightforward that in addition to data, facilities are in an excellent position to provide services (e.g. analysis workflows), a fundamental step for the reproducibility of experiments. However, the quantitative increase in costs and demand for human resources derived from making data FAIR and open, and providing tools for reproducible science is not fully perceived by the RIs and their funders. Currently, many of the development efforts are co-funded by the European Commission through projects, but in the years and months to come (as some of the cluster projects approach their end) there will be the need to seriously consider the sustainability of these new products and procedures, for the EOSC to become tangible.



Figure 2. Sources of funding covering the different phases in the development, upgrade, ordinary and extraordinary maintenance, operation and decommissioning.

All the developments and the progress achieved during these years require funding to be maintained, operated and improved, and there where facilities have partially implemented the data policy, also funding for the implementation and further development of services for their users, but even beyond, for the whole community as envisioned in the EOSC.

This vision of the EOSC is for sure one we all share, where any citizen driven by curiosity can have access to data as well as data visualisation and data processing tools. However, in practical terms, making this vision a reality requires resources, both infrastructure (e.g. networks, computing) and personnel, that are limited. While RIs have been providing access free of charge for many years (as part of their mission and therefore included in their budgets), the exchange of values has always been clear, and most RIs select the best users based on international calls and peer review. This is not the model foreseen for the EOSC, where anyone should be able to access data, free at the point of use. While this generally does not represent a problem for data, this is not the case for computing or storage. To make the picture more complicated, the value for a facility in providing access to these resources for a generic EOSC user is not completely clear, and currently there is no mandate, nor assigned budget, for these purposes. Hence, facilities will need to find alternatives to make their capacity available, but with some kind of cost recovery mechanism. This model, where the scientific resources are provided for a cost, is very unpopular and in some cases even not feasible, since some public institutions cannot charge other public institutions. Nevertheless, public funding is limited and probably not enough to

cover the costs generated by making science reproducible and services accessible to anyone who wants to use them for free.

To motivate funders, and facility managers, to invest in maintaining PaNOSC outputs for EOSC users and the fruitful collaboration between facilities, it is necessary to illustrate who benefits from these and how. The business model/s proposed here will be presented to the relevant stakeholders for their consideration, with the goal to develop a strategy for the sustainability of the PaN EOSC in the years to come.

4.2 Business models' considerations

The business models will be focused not only on the project outputs, but also on the momentum that this project positively drove and how to maintain the collaboration between PaN facilities pursuing similar objectives.

In order to properly develop the business models, considering all the necessary aspects and exchanges involved, it was proposed during the project definition to adopt the business model canvas, a widely known and commonly used tool helping to visualise in a concise way the most important information. The business model canvas was developed mostly for business, as its name indicates, but considering the definition produced by Osterwalder et al: "A business model describes how an organisation creates, delivers, and captures value⁶, in economic, social, cultural or other contexts". The standard business model canvas however, may not be the most appropriate tool to fully articulate the PaN EOSC value proposition. When applying the standard business model canvas, the "client" or "customer" or better "user" in our case, is one, and all the canvas focuses on that single category of users. This approach is fully compatible with the outputs that may have a commercial interest, and is applied to those. But in the case when the exchange involves mostly societal values, while we intend to show the value of these outputs for researchers -our main target-, we also want to make clear that there is a value for the RIs in adopting these outputs. This dual character does not fit into the classic canvas. Therefore, based on the vast experience of our partner EGI with business models for platforms, we adopted a canvas that can consider a multi-sided value proposition, and then adapted the canvas for a multi-stakeholder model.

The PaN EOSC will benefit many stakeholders, including commercial ones, but its main objective is not for profit. If at any step a service will be offered against payment, this will be to cover costs that are otherwise not covered by funders, and not to make profit. The value of the PaN EOSC for other stakeholders is not reflected in the canvas for simplicity, however it will be briefly mentioned.

4.2.1 The business model canvas

The canvas used to represent the business models is the one proposed by Osterwalder in 2010.

4.2.2 The multi-sided value proposition canvas

This canvas has been proposed by Paul Belleflamme and Nicolas Neysen to describe the multi-stakeholder relationships for digital platforms. "Digital platform business models rely on a particular mode of value creation—focused on matchmaking in a multi-sided environment—which highly differs from the traditional vertically integrated models (or 'value chains')."

Although not all of the aspects of the platforms will apply to PaNOSC, one of the features that is relevant in our case is the value of co-creation. In the case of PaNOSC and its outputs, the value is not provided only by

⁶ *Business Model Generation*, [Alexander Osterwalder](#), [Yves Pigneur](#), Alan Smith, and 470 practitioners from 45 countries, self-published, 2010

the facilities but also by users through co-creation, for example for the design of new tools and workflows. In addition, since some of the outputs are based on open source software, value will also be provided by these communities that maintain and improve the open source software.

The goal of using the multi-sided value proposition canvas relies on the need to express a value proposition that addresses the role of the facilities as stakeholders, in addition to the users.

5 Business models

The PaNOSC partners are all different institutions at different states of maturity (some being constructed or just starting user operation while others have over 30 years of experiments done), different sizes and funding models, therefore not all the outcomes of PaNOSC will be adopted by all the partners immediately and there isn't a single business model that fits all.

We will provide a value proposition canvas assuming that all the outcomes are adopted, knowing that this is a simplification.

5.1 Business models for researchers

This is the main scenario and is an evolution of the current situation where RIs users get access to software, hardware and staff support for their experiments. In this scenario we focus on users that have not made an experiment at the RI but that nevertheless want to find and use the data (as per the FAIR principles).

5.1.1 Value proposition

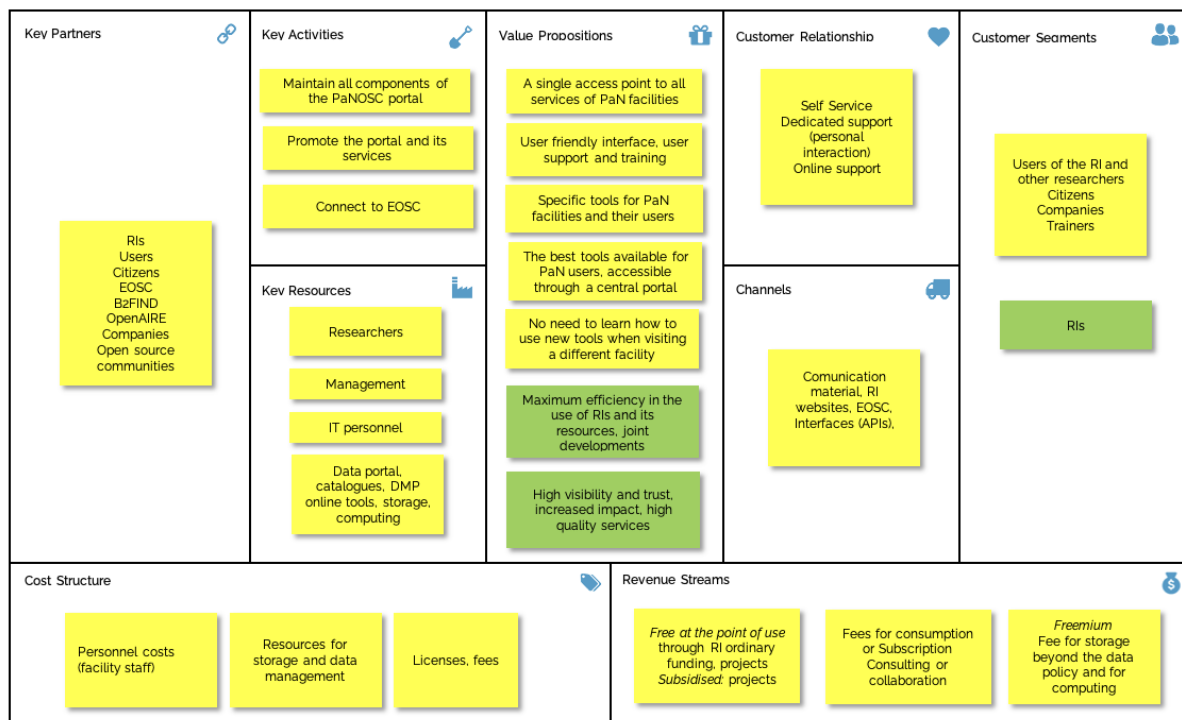
The value proposition⁷ of PaNOSC will benefit multiple stakeholders but the models are focused on the value for the main ones: users and facilities.

The outputs of the project were designed to add value to the user experience but they also add value to the facilities by streamlining and facilitating the delivery of services to users.

⁷ Value proposition is a statement that summarises why a customer should buy a product or use a service

User	PaNOSC outcomes	Facility
<div>Wants</div> <div><ul style="list-style-type: none">To complete their experiments seamlesslyTo perform their experiment, manage and analyse data, and cite relevant datasets in publicationsLess bureaucratic barriersEasy to use, find and access: data and services from all PaN are findable and accessible with a common tool, sharing common standardsTo rely on the IT experts of the facilitiesSimulation to design the experiment</div>	<div>Experience</div> <div><ul style="list-style-type: none">User drivenCommon visuals, common toolsSingle access point to data and services from all/most PaN facilitiesA uniform approach, with shared standardsData commonsCommunity building and co-creationSupport</div>	<div>Wants</div> <div><ul style="list-style-type: none">Provide a high quality service to usersMaximise the outputs of researchAchieve a greater impactMake data and services widely availableAchieve high visibilityAttract new users (the best scientists)Technology benchmark with other facilitiesBetter outreach to industryGood reputation</div>
<div>Needs</div> <div><ul style="list-style-type: none">Reliable and robust servicesData management and data analysis toolsPermanent object identifiersTrusted repositoriesTraining and Help deskIntuitive, user friendly portal and servicesSometimes, free servicesTailored products and services</div>	<div>Benefits & Features</div> <div><ul style="list-style-type: none">Feedback and usage statisticsAdvanced data search tools, specific for PaN dataSoftware catalogue with PaN toolsTraining portalSearch across facilitiesTrusted repositories, some certified (core trust seal)Increased visibility of data and services</div>	<div>Needs</div> <div><ul style="list-style-type: none">Comply with requirements on open data and open scienceMaintain operation costs containedProvide essential services for data management and data analysisRemain competitive with similar RIsMeasure performance and show impact</div>
<div>Fears</div> <div><ul style="list-style-type: none">The new data management practices will add bureaucracyusing different facilities will require to learn multiple proceduresNew services will come at a costSubstantial part of their time will go into data managementData will be leaked before the embargo period is reached.</div>	<div>Trust levers</div> <div><ul style="list-style-type: none">Guaranteed data management quality standardsSoftware rated by usersUsers trusted by facilities</div>	<div>Fears</div> <div><ul style="list-style-type: none">All new procedures and services will require maintenance, increasing facilities' operation costsSubstantial part of resources will need to be available for trainingProviding services to EOSC users will require resources that are not availableProviding services to EOSC users will increase the complexity of servicesInvestment in data storage & resources may not increase the outputs or impact of researchNot enough funds to deliverLarge IT team required</div>
<div>Substitutes</div> <div><ul style="list-style-type: none">Discipline specific portalsEOSC portalCommercial software</div>	<div>Mitigation Actions</div> <div><div><div>USERS</div><div>Threat of substitutes: High</div><div><ul style="list-style-type: none">Specific services (e.g. search tools)Data analysis on site, where the data are generated.</div></div><div><div>FACILITIES</div><div>Threat of substitutes: Medium</div><div><ul style="list-style-type: none">Statistics on user accessCost-benefit analysis</div></div></div>	<div>Substitutes</div> <div><ul style="list-style-type: none">EOSC portalCommercial solutions</div>

5.1.2 Business Model Canvas



This Business Model Canvas shows the following revenue streams and cost structure:

REVENUE STREAMS:

- Freemium
 - Includes both free and paid services. Free services are such only for a limited period or with a basic set of features. If customers want to unlock the full capabilities, they must pay.
 - For data analysis tools a limited capacity could be offered to match the computing and IT resources available in the PaN RIs, potentially via a competitive call for proposals (PaN RIs are only funded and mandated to provide support to users performing experiments). Non-approved proposals and/or requests for higher computing resources than expected would require a payment in the Freemium model.
 - This model, as the pay per use, is not usual and probably would not be well received neither by users nor by facilities, however it is one of the possibilities to make these services sustainable.
- Pay as you go
 - In this business model, users are charged as per the usage of the PaN service in order to recover the full cost of providing these services.
- Consulting or collaboration
 - PaN partners could provide consulting services or conclude a collaboration agreement with users in need to personalise or develop further existing tools. Facilities could charge the client or collaborator on an hourly basis or agree on a budget to complete the activities. In this case, the funds could come from the other entity's budget or from external sources (e.g. projects with external funding).
- Subsidised access by Grants
 - This model allows the user to get services for free, based on a cost recovery arrangement that

is made between the facilities and an entity that grants funding for a limited scope and time.

- Subscription
 - This model allows the user to get services by paying a fixed amount every month or year. In this case, PaN facilities would need to provide enough value for users to benefit from the services recurrently. This model allows to offer a specific number of items under different plans (tiered offering)
- Free at the point of use
 - This model allows the user to get services for free, simplifying the bureaucracy and in some cases, making it accessible to those who could not pay a fee, e.g. due to the regulatory framework of some public institutions. However, this model foresees that all the costs are covered by the entity providing the services, who at the same time needs funders (be it national or EOSC) to be aware and provide dedicated funding.

Please note that some PaN RIs have limits on the amount of commercial revenue they can receive due to their legal structure, therefore extra public funding (via a subscription, new type of associated partnership, new grants) is scalable, while new private funding it is not.

COST STRUCTURE:

In addition to the costs associated with maintaining and adopting the outputs of the project, there is one activity, marketing and communication, that is transversal and necessary to maximise the impact of the PaNOSC offer.

The costs have been categorised as *Annual personnel costs* (recurrent); *Other costs* (these are costs related to each specific activity, such as licensing or subscription to specific services from third parties) and *Other initial costs* (costs that are not recurrent, but have to be sustained once to adopt the outputs of PaNOSC).

- FAIR-compliant Data Policy Framework for Research Infrastructures (KER 1)
 - In order to make the common data policy a reality, it is necessary to adopt/align Data Policies at all PaNOSC sites, and continuously harmonise them. The alignment of policies at PaNOSC partners has been taking place during the project, but the technical implementation of the data policy for some facilities is still in progress, at a different speed in different facilities.
 - The costs related to this activity are highly dependent on the facility, however this is accepted now as a best and required practice and it is expected that the PaN RIs will cover these costs.
- Federated Search API + data portal (KER 2)
 - Creation and maintenance of ontologies, keywords and controlled vocabularies and the data curation related to the catalogue will be necessary after the end of the project. The search tools will need to remain operational at each facility and at the federated Portal/s.
 - Some of the costs related to this activity are accepted as best practice, however keeping the Federated Search API running (and keep updating it to ensure it remains useful and attractive)
- VISA+ web viewers for HDF5 (H5Web, H5Nuvola) (KER 3)
 - VISA is a remote data analysis platform (remote desktop and JupyterLab server).
 - The costs necessary to maintain these outputs refer mostly to personnel costs to continue the collaboration to develop the software and also the costs at each RI to maintain the VISA instances running. The HDF5 viewers follow a similar logic.
- Community AAI (KER 4)
 - The UmbrellaID community AAI (Authentication-Authorisation-Infrastructure) is based on the AARC blueprint architecture and ready for the EOSC AAI federation. PaNOSC is working very closely with GEANT to ensure that UmbrellaID will be one of the first community AAI to form the EOSC AAI federation.
 - The implementation requires the subscription of a Service Level Agreement (SLA) for the

UmbrellaID migration to eduTEAMS (via GÉANT). In addition to these, the maintenance of the UmbrellaID requires mostly personnel costs but could a support contract with GÉANT.

- Simulation package (KER 6)
 - The simulation package VINYL includes SIMEX, McSTAS and OASYS.
 - To maintain the value of the package it is necessary to invest in maintenance and update according to the updates of the open source software. In addition to the maintenance and update, to make the most of VINYL it would be necessary to invest in development and user interface.
- Data Transfer
 - To enable the possibility of downloading large volumes of data or data analysis being performed not in the facility where the data is stored a solution for data transfer is required. In order to support and maintain that solution (e.g. Globus) personnel and license/subscription fees would be required.
- Community e-learning and training platform (KER 5)
 - The training catalogue and e-learning platform (delivered jointly with the ExPaNDS project) will require maintenance, curating of content and hosting. It will also be dependent on other outputs from PaNOSC (VINYL, VISA, etc.)

5.2 Business models for commercial users

For the purposes of this section commercial users must be understood as any other user that can profit from the tools and services provided (e.g. a private company doing research on their field of interest).

5.2.1 Business model I: Complete environment for data search, analysis, simulation and storage

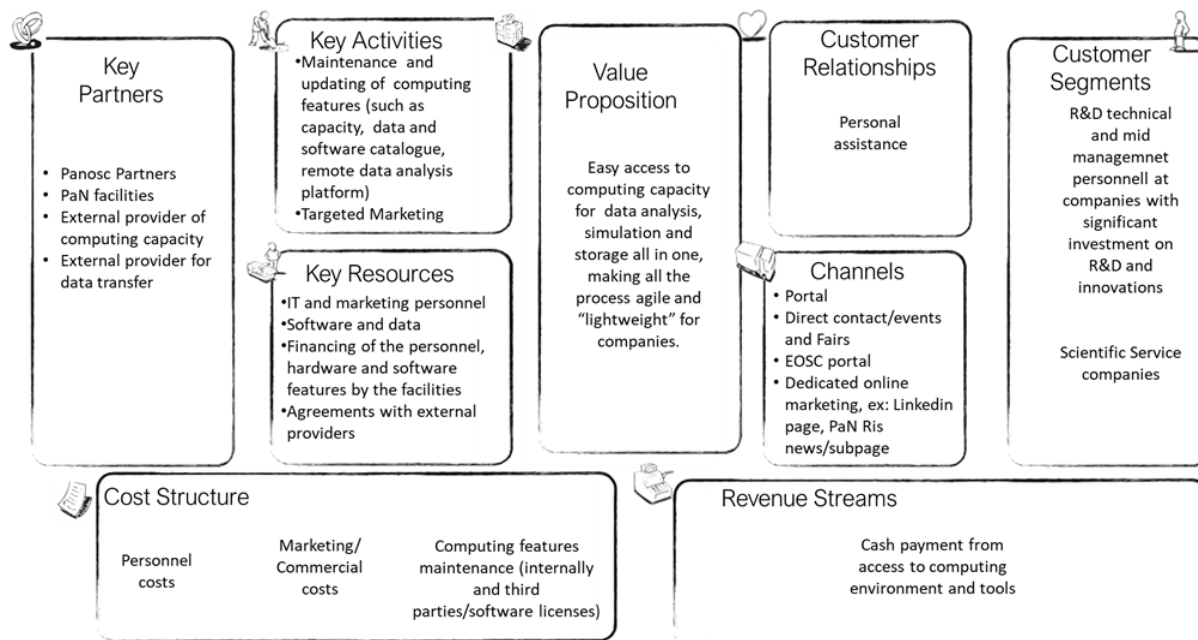
This business model is quite similar to the current interactions with commercial users that use some of the RIs, the difference being that rather than making their own experiment and working with their own data this time the commercial user will not perform an experiment and just use existing open data.

5.2.1.1 Value proposition

The value proposition of PaNOSC will be similar to the one shown for Researchers above:

Commercial User		PaNOSC outcomes		Facility	
Wants		>	Experience	<	Wants
<ul style="list-style-type: none">• To complete their experiments seamlessly• To manage and analyse data• Easy to use, find and access: data and services from all PaN are findable and accessible with a common tool, sharing common standards• To rely on the IT experts of the facilities• Simulation to design the experiment			<ul style="list-style-type: none">• User driven• Common visuals, common tools• Single access point to data and services from all/most PaN facilities• A uniform approach, with shared standards• Data commons• Support		<ul style="list-style-type: none">• Provide a high-quality service to users• Maximise the outputs of research• Achieve a greater impact• Increase cooperation with commercial partners (some RIs are mandated to do this)• Technology benchmark with other facilities• Better outreach to industry• Good reputation
Needs		>	Benefits & Features	<	Needs
<ul style="list-style-type: none">• Reliable and robust services• Data management and data analysis tools• Permanent object identifiers• Trusted repositories• Training and Help desk• Intuitive, user friendly portal and services• Tailored products and services			<ul style="list-style-type: none">• Feedback and usage statistics• Advanced data search tools, specific for PaN data• Software catalogue with PaN tools• Training portal• Search across facilities• Trusted repositories, some certified (core trust seal)• Increased visibility of data and services		<ul style="list-style-type: none">• Comply with requirements on open data and open science• Maintain operation costs contained / recover the costs of the extra services• Provide essential services for data management and data analysis• Remain competitive with similar RIs• Measure performance and show impact
Fears		>	Trust levers	<	Fears
<ul style="list-style-type: none">• Access to the data and services is slow and bureaucratic• Using different facilities will require to learn multiple procedures• New services will come at a high cost• Data will be leaked before the embargo period is reached.			<ul style="list-style-type: none">• Guaranteed data management quality standards• Software rated by users• Users trusted by facilities• Tailored service for paying commercial customers		<ul style="list-style-type: none">• All new procedures and services will require maintenance, increasing facilities' operation costs• Substantial part of resources will need to be available for training• Opportunity cost of servicing commercial users vs. researchers• Investment in data storage & resources may not increase the outputs or impact of research• Not enough funds to deliver• Large IT team required
Substitutes		>	Mitigation Actions		Substitutes
<ul style="list-style-type: none">• Discipline specific portals• EOSC portal• Commercial software			<u>USERS</u> Threat of substitutes: High <ul style="list-style-type: none">• Specific services (e.g. search tools)• Data analysis on site, where the data are generated.	<u>FACILITIES</u> Threat of substitutes: Medium <ul style="list-style-type: none">• Statistics on user access• Cost-benefit analysis	<ul style="list-style-type: none">• EOSC portal• Commercial solutions

5.1.2.2 Business Model Canvas



This Business Model Canvas shows the following revenue streams and cost structure:

REVENUE STREAMS:

- The revenue will come in form of cash deriving from selling access to the computing environment and all the tools for data search, data analysis and simulation. The access will be sold through an agreement that could imply the access for a limited time window (monthly or annual), under the payment of a fee, that can be renovated.

Please note that some PaN RIs have limits on the amount of commercial revenue they can receive due to their legal structure.

COST STRUCTURE:

- The cost structure will be the same as for Researchers given that this business model assumes a complete package

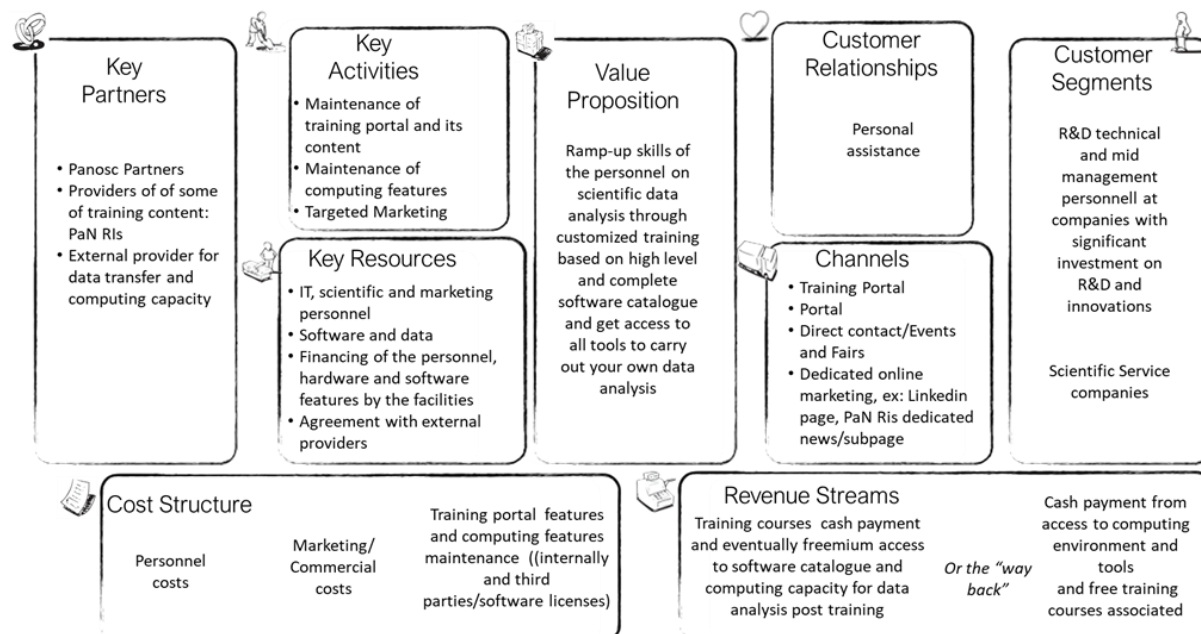
5.2.2 Business model II: Improving company's staff skills on data analysis and access to related tools

This business model focuses on partnering with commercial users in order to train and support their staff in data analysis.

5.2.2.1 Value proposition

Commercial User		PaNOSC outcomes		Facility		
Wants		>	Experience	<	Wants	
<ul style="list-style-type: none">To learn data analysis best practicesTo learn how to benefit from open dataTo train their staffTo manage and analyse dataEasy to use, find and access: data and services from all PaN are findable and accessible with a common tool, sharing common standards			<ul style="list-style-type: none">User drivenCommon visuals, common toolsSingle access point to data and services from all/most PaN facilitiesData commonsSupport		<ul style="list-style-type: none">Provide a high-quality service to commercial usersIncrease cooperation with commercial partners (some RIs are mandated to do this)Better outreach to industryGood reputation	
Needs		>	Benefits & Features	<	Needs	
<ul style="list-style-type: none">Staff trained in data management and data analysis toolsHelp deskData management and data analysis toolsIntuitive, user friendly portal and servicesTailored products and services			<ul style="list-style-type: none">FeedbackTraining portalTrusted repositories, some certified (core trust seal)Increased visibility of data and services		<ul style="list-style-type: none">Establish good working relationships with commercial usersProvide essential services for data management and data analysisRemain competitive with similar RIsMeasure performance and show impact	
Fears		>	Trust levers	<	Fears	
<ul style="list-style-type: none">Access to the data and services is slow and difficultUsing different facilities will require to learn multiple proceduresNew services will come at a high costData will be leaked before the embargo period is reachedStaff new skills will not be worth the investment			<ul style="list-style-type: none">Guaranteed data management quality standardsTraining and Software rated by usersTailored service		<ul style="list-style-type: none">Substantial part of resources will need to be available for trainingOpportunity cost of servicing commercial users vs. researchersInvestment in training may not increase the outputs or impact of researchNot enough funds to deliverLarge IT team requiredLarge scientific team required / disruption of the scientists research work	
Substitutes		>	Mitigation Actions		<	Substitutes
<ul style="list-style-type: none">Discipline specific portalsEOSC portalCommercial software			<u>USERS</u> Threat of substitutes: High <ul style="list-style-type: none">Specific services (e.g. search tools)Data analysis on site, where the data are generated.	<u>FACILITIES</u> Threat of substitutes: Medium <ul style="list-style-type: none">Statistics on user accessCost-benefit analysis		<ul style="list-style-type: none">EOSC portalCommercial solutions

5.2.2.2 Business Model Canvas



This Business Model Canvas shows the following revenue streams and cost structure:

REVENUE STREAMS:

- The revenue from the whole value offered can be collected in two ways. It has been considered to:
 - a) Obtaining cash payment from providing access to training courses and eventually associate a freemium access to software catalogue and computing capacity for data analysis post training.

Or

- b) Obtaining a cash payment from access to computing environment and tools and offering free training courses associated.

In both cases, being the value of this business model higher than the one offered in business model I, the price of this offer will be established accordingly.

COST STRUCTURE:

- In this model the costs categories will be the same as in Business Model I, but with a higher cost on scientific personnel, related to the analysis of data and reporting, that will be estimated.

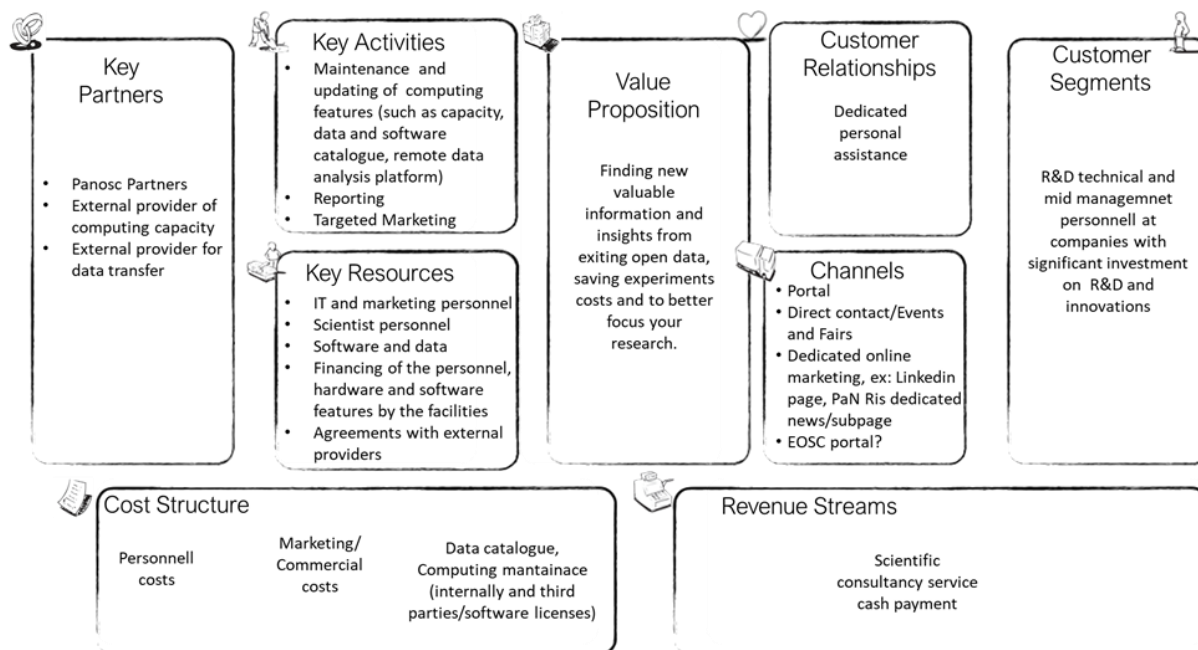
5.2.3 Business model III: Complete consultancy service

This business model focuses in offering consultancy services to private entities, for example understand which kind of results with photon and neutron facilities have been obtained on certain materials and eventually make the data re-analysed by facilities personnel with different software in order to find new valuable information and insights from exiting open data. This can lead not only to saving experiments costs for companies but also to better focus their research.

5.2.3.1 Value proposition

Commercial User		PaNOSC outcomes		Facility		
<div>Wants</div> <ul style="list-style-type: none">To learn new insights based on existing dataTo benefit from state-of-the-art research facilities without owning the long-term associated costsTo manage and analyse dataQuick and easy to use service	>	<div>Experience</div> <ul style="list-style-type: none">User drivenPersonalised Support	<	<div>Wants</div> <ul style="list-style-type: none">Provide a high-quality service to commercial usersIncrease cooperation with commercial partners (some RIs are mandated to do this)Better outreach to industryGood reputation		
<div>Needs</div> <ul style="list-style-type: none">Research done on a certain topicData management and data analysis toolsIntuitive, user friendly portal and servicesTailored products and services	>	<div>Benefits & Features</div> <ul style="list-style-type: none">FeedbackTrusted repositories, some certified (core trust seal)Increased visibility of data and services	<	<div>Needs</div> <ul style="list-style-type: none">Establish good working relationships with commercial usersRemain competitive with similar RIsMeasure performance and show impact		
<div>Fears</div> <ul style="list-style-type: none">Access to the research insight is slow and difficultUsing different facilities will require to learn/pay multiple proceduresNew research will come at a high costConsultancy service will not be worth the investment	>	<div>Trust levers</div> <ul style="list-style-type: none">Guaranteed data management quality standardsTraining and Software rated by usersExperience in researchTailored service	<	<div>Fears</div> <ul style="list-style-type: none">Substantial part of resources will need to be available for consultancyOpportunity cost of servicing commercial users vs. researchersInvestment in consultancy may not increase the outputs or impact of researchLarge teams required		
<div>Substitutes</div> <ul style="list-style-type: none">Commercial ConsultancyEOSC portalPartnership with universitiesOwn research team	>	<div>Mitigation Actions</div> <table><tr><td><div>USERS</div><div>Threat of substitutes: High</div><ul style="list-style-type: none">Specific services (e.g. search tools)Data analysis on site, where the data are generated.</td><td><div>FACILITIES</div><div>Threat of substitutes: Medium</div><ul style="list-style-type: none">Statistics on user accessCost-benefit analysis</td></tr></table>	<div>USERS</div> <div>Threat of substitutes: High</div> <ul style="list-style-type: none">Specific services (e.g. search tools)Data analysis on site, where the data are generated.	<div>FACILITIES</div> <div>Threat of substitutes: Medium</div> <ul style="list-style-type: none">Statistics on user accessCost-benefit analysis	<	<div>Substitutes</div> <ul style="list-style-type: none">EOSC portalCommercial solutions
<div>USERS</div> <div>Threat of substitutes: High</div> <ul style="list-style-type: none">Specific services (e.g. search tools)Data analysis on site, where the data are generated.	<div>FACILITIES</div> <div>Threat of substitutes: Medium</div> <ul style="list-style-type: none">Statistics on user accessCost-benefit analysis					

5.2.2.3 Business Model Canvas



This Business Model Canvas shows the following revenue streams and cost structure:

REVENUE STREAMS:

- The revenue will come in form of cash deriving from selling the consultancy through a service agreement.

COST STRUCTURE:

- In this model the costs categories will be the same as in Business Model I, but with a higher cost on scientific personnel, related to the analysis of data and reporting, that will be estimated.