1 Overview

The members of the Scientific Advisory Board[1] of the Helmholtz Incubator Platform „Helmholtz Federated IT Services" (HIFIS) gathered for their (virtual) second meeting in June 2021.

In the meeting, the progress of HIFIS has been presented and discussed. Presentation materials[2] as well as the annual HIFIS report of 2020[3] have been provided to all members. Last year's report[4] of the SAB was available for reference and comparison. Minutes of the meeting, including responses to questions raised, were accessible during and after the conference[5].

Adding to the fruitful discussions during the meeting, the SAB members were asked to provide written feedback on issues of specific interest to them. The full feedback from several SAB experts is appended in the following chapters. In the following, we briefly summarize the overall feedback of the SAB.

1.1 Summary

The feedback from SAB experts generally acknowledged a very positive, "first-rate" evolution of HIFIS, as seen by the "very good progress in terms of the infrastructure and services offered to the users". Further, it was stated that the "increase on number of users is impressive". A overarching advantage "is undoubtedly the establishment of this team spread over several centers and the pooling of their know-how".

It was noted that it "will be important to have clear definitions of the success factors for the HIFIS. These will be technical, political and social; and, how they will be measured from the perspective of the users of HIFIS." Some discussion arose amongst SAB members with respect on the proper weighting of success factors. Several advisers focused on close collaboration in high level/widely used scientific (software) projects. HIFIS "should focus on enhancing the use cases that will emerge from the communities themselves". In coordination with other platforms (e.g., HMC); it is encouraged to tackle to problem of distributed data, such as a "superordinate data directory" and to "include metadata concepts that will enable the scientific community to better identify its data in the long term". In another vote, these partly concentrate on aspects "that are outside HIFIS itself": "HIFIS mainly provides infrastructure that is only partially used by the scientific community directly. These areas access HIFIS services as a central component and provide services for their communities."

---

[1] https://hifis.net/sab
[2] https://events.hifis.net/event/136/timetable/#20210614.detailed
[3] https://nubes.helmholtz-berlin.de/s/Y79kf2s3xyiZnX2/download
[4] https://nubes.helmholtz-berlin.de/s/ffK2iAgiDqFLgGJ?path=%2F&openfile=84005457#pdfviewer
[5] https://notes.desy.de/F6wjgM0tTdeFrsxKwX9sNA?view#

Concerning the current service portfolio, it was stated that "the services opened so far are a good contribution to the support of scientists." With respect to the scope and number of trainings, "the set of services provided so far is a good selection". The education programme is encouraged to "enlarge this in the next years", for example on "actual usage of computing or memory of the applications run in the cloud", presumably to increasingly "support scientists to be successful".

Concerning the further development of the cloud services portfolio, it was stressed that "expressions of needs from researchers in terms of data processing (experimental or from numerical simulations) should now be collected". Based on this, "indications on how to make HIFIS interact with national or international e-infrastructure (whether thematic or generalist) could thus be specified".

The SAB members accentuated the sustainability of HIFIS and its services: "HIFIS success needs to be confirmed in the future certainly by ensuring the sustainability of current services (including their upgrades). It would also be useful to set up a process to manage their lifecycle". In that context, it was discussed and advised to "plan which is the number of services that you can support in order not to surpass your capacity in a way that adequate service cannot be provided".

A further issue is the stronger focus on security topics opened by the newly established cloud services. "An overall view of cyber security for HIFIS service should be considered". HIFIS might have a more central position "to define, audit and manage overall security".

With respect to the integration of the board to HIFIS work and progress, at least one advisor indicated the wish for a different setup and "would have liked to see a different form of organization for the advisory board, one that actually enables the board members to meet and discuss internally, in the sense of the term 'board'."

With the (pilot version of the) Helmholtz Cloud being already "usable and well-known", it is advised that this HIFIS-led integration "must be further expanded so that cooperation partners can be integrated". Closer collaboration and even consulting with international initiatives, such as EOSC, is strongly encouraged. It was made clear that "the experience acquired in building such a global set of services may profit to other countries/organizations".

2   Feedback of Expert 1

2.1   Service Portfolio Sustainability

We have had two meetings which have been very focused on details of the HIFIS. This is very insightful especially at this development phase of HIFIS. It may helpful to the SAB, as part of subsequent meetings, to stand back and consider more strategic aspect of HIFIS, such as the cost and resource model for HIFIS services; both for development and operations. In doing so it is important to establish that the service portfolio is sustainable.

2.2   Definition of Success factors

It will be important to have clear definitions of the success factors for the HIFIS. These will be technical, political and social; and, how they will be measured from the perspective of the users of HIFIS. This will help inform the value of services but also inform on gong investment in HIFIS.

2.3   Cyber Security

In response to question on management of cyber security it was explained to be the responsibility of service providers in each of the Helmholtz centres. However an overall view of cyber security for HIFIS service should be considered. To define, audit and manage overall security of HIFIS. Without which the whole of the federated HIFIS service could be vulnerable to some local weakness.

3   Feedback of Expert 2

My feedback with regard the evolution of HIFIS is very positive. I have seen very good progress in terms of the infrastructure and services offered to the users. The increase on number of users is impressive.

Also, the set of services provided to the users looks good to me. The portal is also a very convenient point of entry for the users. Processes for support, consultancy and training look good also.

I recommend to make sure you can cope with a continuous increase of users in the near future. Also, it is important to plan which is the number of services that you can support in order not to surpass your capacity in a way that adequate service cannot be provided.

The current training offering seems a bit short and limited to project management and programming in Python. It would be important to design a good program that gives answers to the user community.

The contribution of HIFIS to the German and European landscape is clear to me. I can only recommend you to continue with the good progress.

I am answering below some of the questions:

- What are key points indicating the success of HIFIS?

I think that clearly the progress in the number of services and users that are now present in HIFIS is a measure of the success. I can only congratulate you for this.

- How can Helmholtz Cloud and Backbone Services support scientists to be successful?

The services opened so far are a good contribution to the support of scientists. However, it is a bit surprising that in the survey a large number of your users declare that they were not using any of the provided services.

An aspect that I saw a bit short is the number of trainings. I guess you can enlarge this in the next years. Maybe you can provide more trainings about the services in order to increase the number of users.

- How to balance basic and advanced topics in RSE education?

This is always a tricky question. An option is to ask the users for the courses that they would like to see, or use surveys to classify the topics. An approach given that nowadays most courses are online is to record the courses and offer some of them just as recordings. This allows to extend the number of available courses without increasing the effort to deliver them.

- What else could Software Services do to support scientists to be successful?

The set of services provided so far is a good selection. Some services that can be added can be related to the detailed performance behaviour of the applications run in the cloud. I am not sure if this is already available, I browsed the services catalogue and could not find it. For example, information about the actual usage of computing or memory of the applications run in the cloud, to enable the users to decide if they should use a larger number of nodes or a larger instance for their executions.

- How to fairly distribute consulting?

My approach would be to distribute it proportional to the demand. Or also, proportional to the number of users of each centre.

4   Feedback of Expert 3

From the report and the presentations made during the meeting, it is quite clear that the work carried out to set up HIFIS services is first-rate. In particular, and without

underestimating the technical difficulties that had to be faced, one of the main successes is undoubtedly the establishment of this team spread over several centers and the pooling of their know-how.

As I understand it, the cloud services deployed were selected by means of surveys. This is a good starting point. However, in order to ensure a wider adherence to the use of HIFIS, it seems to me that expressions of needs from researchers in terms of data processing (experimental or from numerical simulations) should now be collected. Valuable indications on how to make HIFIS interact with national or international e-infrastructure (whether thematic or generalist) could thus be specified.

The HIFIS ecosystem is rich and requires significant resources. The human resources required are numerous (43 FTE). It is not clear to me what is the distribution between permanent and temporary staff among these FTE. Depending of it, the sustainability of such or such services may be at risk.

Furthermore HIFIS success needs to be confirmed in the future certainly by ensuring the sustainability of current services (including their upgrades). It would also be useful to set up a process to manage their lifecycle (when, why and how a service may be decommissioned or needs to be expanded,...)

The experience acquired in building such a global set of services may profit to other countries/organizations: is it foreseen doing some consulting on this ?

5 Expert 4

Feedback on the 2nd Advisory Board Meeting

5.1 Preliminary remarks

It is very difficult for an outsider to visualize all the facts about all the elements of the Helmholtz Cloud. In fact, it is rather the case that despite studying the documents and attending the conference, far too few details can be fully remembered in relation to the whole. I write this as an apology in case my comments are inaccurate or my suggestions have already been implemented. In that case, please take my statements as confirmation and approval.

I would have liked to see a different form of organization for the advisory board, one that actually enables the board members to meet and discuss internally, in the sense of the term "board". The consideration of one's own comments, a joint opinion, etc. are central elements that, together with a joint report, enable a more sustainable discussion. The organizational form chosen by HIFIS leaves lonely reviewers:

- You can't get an answer to even the simplest question, "Say, how did you understand that?"

- Sentences such as I am writing now must be written by everyone.
- Not knowing at first what the tenor of the last evaluation was, I felt this session lacked at least a general benchmark for assessing progress and improvement. Perhaps the summary of board member feedback could be communicated in a more open and timely manner.

5.2 Summary

According to the descriptions, the HIFIS project has made very good progress over the past year. According to the developments described, pushing the implementation of the underlying, open identity management in particular has paved the way for the use of the teaser applications gitlab and mattermost, thus making the Helmholtz Cloud usable and well-known. This must be further expanded so that cooperation partners can be integrated as a matter of course.

5.3 Focus of my comments

I could not identify with some of the comments and questions during the conference. In my view, they were aimed too much at the areas of the Helmholtz Cloud that are outside HIFIS itself. My understanding is that HIFIS mainly provides infrastructure that is only partially used by the scientific community directly. These areas access HIFIS services as a central component and provide services for their communities. In my view, this is especially true for the required user-specific training and the additional KPIs that this requires. Aligning KPIs with the number of using centers and the number of using services, such as for AAI, are absolutely sufficient for now, at least in my opinion. Guidance for HIFIS should therefore be aimed at the providers of these services rather than the actual users - this is then also the direction of my comments below.

5.4 Application support through instructions

The focus on supporting the use and application of gitlab by focusing on the DevOps area and developing suitable workflows should serve as an example for supporting all applications. In terms of the somewhat confusing service offering, guidance on how to use these services could remove further barriers to entry when using them. This is especially true for the HIFIS applications that are also directly in contact with science.

Further guidance should be provided for the use of HIFIS infrastructure services, e.g. for the use of compute clusters and storage services - compute services as required for CI/CD or for the realisation of large scientific applications, for example, and storage services that do not simply offer sync&share or S3-compatible access, but also take care of the data itself. (Data Management). Whereby by instructions I mean not only explanations of the basic use but also Hifis-specific instructions or guidelines for the use of the services (—> data management, acceptable use, ...).

5.5   Focus on data management as an infrastructure task

5.5.1 Data directory for users, communities, topics

From experience, the use of cloud services leads to an unfavorable distribution of documents or, more generally, of data of the users or their communities in the long run. Data is usually stored close to the cloud services and is very difficult for individuals to keep together. Several unclearly delimited services with comparable or even identical functions, several accounts of one person, for example because of different roles in several collaborations, are examples that show how this confusion can be amplified. In HIFIS, both are given.

From my point of view, a superordinate data directory is missing, which can support the following, among other things:

- the sustainability of long-term use
- IT security with regard to availability (backup)
- compliance with the rules of good scientific practice (archiving, versioning, ...)
- the possibility of locating certain data in order to comply with data protection requirements

5.5.2 Instructions with metadata concepts

The following would therefore be advisable:

- The aforementioned guidance should be expanded to include metadata concepts that will enable the scientific community to better identify its data in the long term.
- Infrastructures could be encouraged and enabled to integrate these concepts into their data services.
- The applications offered/used could be further developed in this respect, or their further development could be encouraged or even promoted.

5.5.3 Identification of and handling with duplicates

These metadata concepts must above all support the identification of duplicates and insist on their attention. The identification of necessary post-synchronizations is only one task. One must also prepare for the sensible handling of data in cloud environments:

- How easily data is given new timestamps when copied or
- copied past the Sync&Share service to a remote infrastructure for use in another cloud application.

Therefore it is also necessary in the infrastructures not to rely only on the metadata in file systems and on the features of Sync&Share solutions.