

Deliverable 2.4

Project ID	654241
Project Title	A comprehensive and standardised e-infrastructure for analysing medical metabolic phenotype data
Project Acronym	PhenoMeNal
Start Date of the Project	1st September 2015
Duration of the Project	36 Months
Work Package Number	2
Work Package Title	Sustainability of PhenoMeNal
Deliverable Title	D2.4 Business plan beyond the end of the project
Delivery Date	M24
Work Package leader	UL
Contributing Partners	UL, UU, EMBL-EBI, IPB Halle, U Jena, UB, CIRMMMP
Authors	Merlijn van Rijswijk, Ola Spjuth, Steffen Neumann, Namrata Kale, Chris Steinbeck, Thomas Hankemeier, Michael van Vliet, Claire o'Donovan, Pablo Moreno, Ralf Weber, Antonio Rosato
Abstract	For the sustainability of PhenoMeNal, we will set up a foundation. The foundation will oversee the further development and outreach to users after the end of the PhenoMeNal project.



1	EXECUTIVE SUMMARY	3
2	DETAILED REPORT OF THE DELIVERABLE	4
2.1	Economic sustainability strategy	4
2.2	PhenoMeNal Foundation	5
2.2.1	<i>Governance structure</i>	5
2.2.2	<i>Sources of Income</i>	6
2.2.3	<i>Heart of the PhenoMeNal community</i>	6
3	TECHNICAL SUSTAINABILITY	6
3.1	Key components	7
3.2	Portal	7
3.3	Infrastructure deployment logic (KubeNow)	7
3.4	Services	7
3.5	Technical sustainability scenarios	8
3.5.1	Funding sources for the scenarios	10
4	METABOLOMICS ELIXIR USE CASE	11
5	IMPLEMENTATION NETWORK GO-FAIR/EUROPEAN OPEN SCIENCE CLOUD	11
7	WORK PLAN	14
8	CONCLUSIONS	14
9	Appendix 1 Current installations of PhenoMeNal	15
9.1	Current installations of PhenoMeNal within the consortium	15
9.1.1	Uppsala University	15
9.1.2	The Netherlands / Leiden University	15
9.1.3	EMBL-EBI	16
9.1.4	Germany (de.NBI)	16
10	Appendix 2 National policy measures supporting the development of PhenoMeNal in the next years	18
10.1.1	UK Institute for Health and Biomedical Informatics Research	18
10.1.2	UKMedBio infrastructure. UKMedBio / Medical Research Council	18
10.1.3	The Netherlands: X-omics initiative	18
10.1.4	Swedish National Bioinformatics Infrastructure Sweden, (http://nbis.se/)	19
10.1.5	Swedish Tryggve (https://wiki.neic.no/wiki/Tryggve)	19



1 EXECUTIVE SUMMARY

For the sustainability of PhenoMeNal, we will set up a foundation and make sure this is active before M30. The foundation will oversee the further development and outreach to users after the end of the PhenoMeNal project. This is of great importance, to give (potential) users, industry and (science) funders trust that PhenoMeNal will be further developed and maintained in a coordinated manner.

To plan of approach to grow a “sticky” user base has been described in D2.3 “Updated Report on mapping of e-infrastructures, users, investments for supporting policy developments in the field of metabolomics, biomarkers and biobanks”, which was submitted two months ago in M22.

The foundation will also steer the European Open Science Cloud Metabolomics Implementation Network. We expect the foundation will be primarily funded by (annual) partnership fees. Likely, the partnership fees will be enough to run the foundation, establish and maintain a website, and to organise an annual “community” meeting for developers and users. However, the income is unlikely to be high enough for the sustained development of PhenoMeNal. For this, we will need to rely on acquisition of grants and subsidies, both on a national and European/H2020 level.

The core groups involved in the development of PhenoMeNal have been active to acquire (limited) national funding for the upkeep of PhenoMeNal according to the “minimal maintenance scenario” in the next 24 months after the end of the project. An overview of the national funding applications/grants is listed in Appendix 2. To determine funding strategies for the more elaborate scenarios “Actively maintained VRE” and “Actively maintained and further developed VRE” we will organise a workshop in early autumn, when the H2020 Infrastructure programme 2018/2019 has been finalised and published. Ideally this would be a similar project in size as PhenoMeNal, now with a broader scope including the integration of proteomics, genomics and other omics data sources, allowing us to carry out the “Actively maintained and further developed VRE” scenario. However, we will also consider coordination and support actions, with a more limited budget, to carry out the “Actively maintained VRE” scenario. In addition, several other initiatives have been taken, among which establishing a Metabolomics Use case in ELIXIR and setting up an Implementation Network in the EOSC. However, these initiatives are likely only to contribute (financially) in the long term to the sustainability of PhenoMeNal.



2 DETAILED REPORT OF THE DELIVERABLE

In D2.2. "A first version of the sustainability plan, which details business models options as well as the design choices and risk mitigation if any of the components fails to be sustainable", submitted in M12, we presented the following (economical) sustainability strategies:

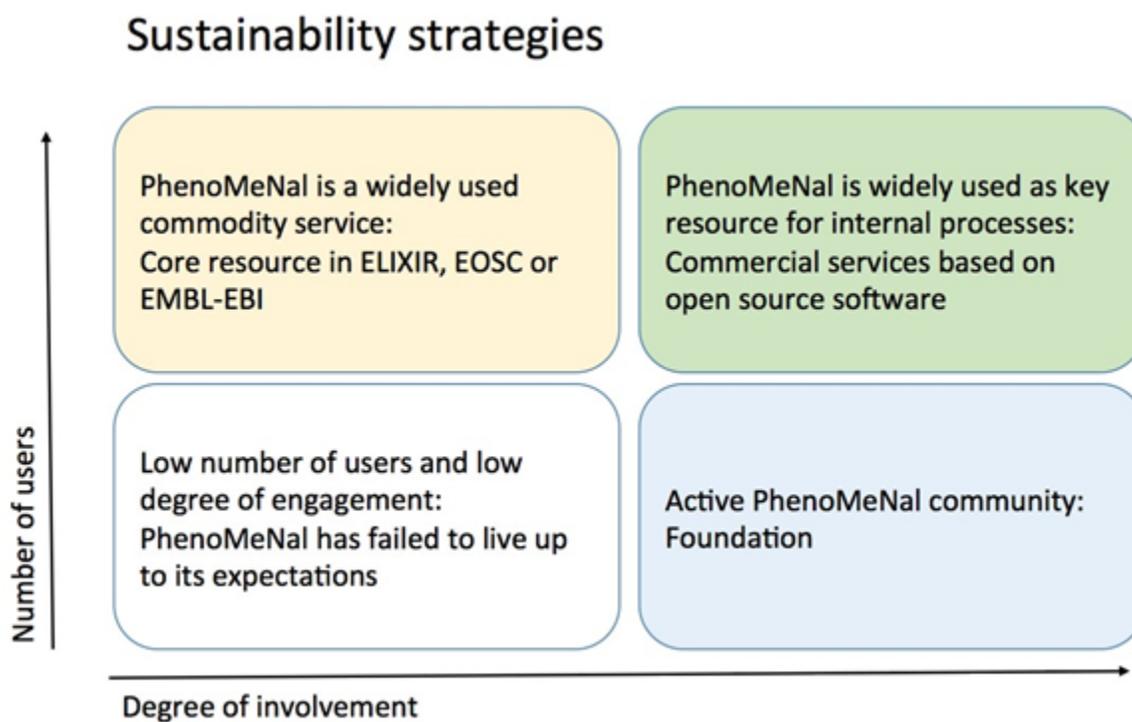


Figure 1. Sustainability strategies for PhenoMeNal.

2.1 Economic sustainability strategy

PhenoMeNal is still in the (technical) build-up phase, end of February 2017 a first beta-release of PhenoMeNal took place, and the first production phase release was at the end of August 2017. This means in practice that only now, having a product in hand, is the time to reach out to users outside of the consortium and to grow a user base. To this end we have presented a "plan of approach" in D2.3 "Updated Report on mapping of e-infrastructures, users, investments for supporting policy developments in the field of metabolomics, biomarkers and biobanks", which was submitted two months ago, in M22.



Given that we have only a limited user base at this moment in time, and highly involved developing partners, our **first choice for the sustainability** of PhenoMeNal is “**Active PhenoMeNal community, foundation**”. Over time, after the end of the project, this could develop when the user base is growing as anticipated, to “Commercial services based on open source software” or eventually even when PhenoMeNal would become a commodity to “Core resource in ELIXIR, EOSC or EMBL-EBI”.

The consequences of this choice are:

1. We will set up a Metabolomics Foundation and make sure this is active before M30. A more elaborate discussion on the scope, activities, business model and governance of the foundation follows later in this document.
2. Growing the user base is of utmost importance. We will focus on “champions”, influential (biological) users whose publications can help in growing the acceptance and user base of PhenoMeNal, and active users who can help continue the development of the PhenoMeNal infrastructure after the end of the project.

2.2 PhenoMeNal Foundation

Currently we are preparing to establish a Foundation, to oversee the further development and outreach to users after the end of the PhenoMeNal project. This is of great importance, to give (potential) users, industry and (science) funders trust that PhenoMeNal will be further developed and maintained in a coordinated manner. We intend to have the foundation active before M30.

We anticipate the Foundation to have a broad mission, in order for it to be usable in the future for other activities as well without a need to change the Articles of Association, for instance to govern the European Open Science Cloud Metabolomics Implementation Network. Its application area will not be limited to metabolomics, but will also include the integration of various omics data and omics data analysis pipelines in general. The activities will also be broadly formulated, to include outreach, communication, networking and research coordination activities in general.

2.2.1 Governance structure

We intend to have a Board consisting of up to seven board members. The Members will represent the various actors: universities developing technology, clinics, industry (including instrument vendors) and users in general applying metabolomics in their research. Besides the board there will be a General Assembly, meeting at least once a year with the board. All partners of the foundation can nominate one member for the General Assembly. The board will be assisted by one or more committees. We



anticipate that one of the committees will be responsible to oversee the further development of PhenoMeNal, and to advise the Board on this when actions are needed.

2.2.2 Sources of Income

We expect the foundation will be primarily funded by (annual) partnership fees. Currently we are testing which level of partnership fees are feasible, Likely for academic organisations a partnership of EUR 1.000 - 2.000 per year is acceptable, for industry we expect a partnership fee of EUR 5.000 - 10.000. These figures indicate that the partnership fees will be enough to run the foundation, establish and maintain a website, and to organise an annual “community” meeting for developers and users, but is likely not enough for the sustained development of PhenoMeNal. For this, we will need to rely on acquisition of grants and subsidies, both on a national and European/H2020 level.

2.2.3 Heart of the PhenoMeNal community

For the remaining time of the project it is crucial to engage and to build an active user and (developer) community around the PhenoMeNal specific components. We need to educate and convince others about the benefits when developing tools and workflows the PhenoMeNal way and applying these in research. To this end we have developed a plan in D2.3 to reach out to users, “champions/ambassadors” (biological and medical researchers influential in their field, who can enable us to reach out to a broader community for accepting and implementing PhenoMeNal.) which we will implement in M25-M36. In M25 we will organise a PhenoMeNal summer school in Sardegna, Italy. The training material developed for this course will be used for further trainings, for instance during the MetaboMeeting in Birmingham in December 2017 and for 1:1 trainings on-site.

3 TECHNICAL SUSTAINABILITY

Technical sustainability has been high on the priority list since day one of the project. The strategy has always been to (re-)use, preferably actively maintained, open source components where possible. All missing components that have been developed in PhenoMeNal are open source and publicly available at GitHub. A full history of both code and documentation is kept for future reference. Additions or improvements to existing open source components have when possible been contributed back to the original maintainers according to their license. In Appendix 1 an overview has been given on the implementation of PhenoMeNal at the various partners of the consortium.



3.1 Key components

PhenoMeNal consists out of several key components. Without these components users will not be able to use our services as designed.

3.2 Portal

The portal provides a user-friendly interface to the PhenoMeNal services. It helps users to decide where to deploy the services to, and will guide, especially non-technical, users when setting up their own private instance of PhenoMeNal. At the time of writing it still hosted in one location making it a bottleneck from a sustainability perspective a single point of failure (SPOF) and is therefore being redeveloped as a portable component that can be made accessible at multiple sites, both public and private. We note that the portal is not strictly required to instantiate VREs, but greatly simplifies the process for users with a graphical user interface.

3.3 Infrastructure deployment logic (KubeNow)

KubeNow is used to instantiate PhenoMeNal VREs. KubeNow is built as a thin layer on top of established tools Terraform, Ansible, and KubeAdm supported by large developer and user communities. KubeNow is in good shape, and after M36 the challenge will be to adapt KubeNow to the moving virtual infrastructure frameworks target. There will be low resources required to keep KubeNow with current functionality, which will suffice to sustain PhenoMeNal, but additional effort will enable PhenoMeNal to take advantage of the latest developments in the field, such as features for improved high-availability and autoscaling.

3.4 Services

1. *Long-running services*
 - a. **Galaxy.** Galaxy and the Galaxy-Kubernetes runner are exciting products. Maintaining current state is not much effort, but the tools require further development in order to provide all the features a user needs when working within the VRE. Hence this is one of the more challenging components of PhenoMeNal from sustainability perspective.
 - b. **Luigi.** Luigi runner for Kubernetes has been pushed upstream to Spotify, and there is very little effort required to sustain Luigi in PhenoMeNal.
 - c. **Jupyter.** Jupyter integration in VRE is simple, and there is very little effort required to sustain Jupyter in PhenoMeNal.
2. *Ephemeral services.* These comprises all analysis components that have been containerised and made Available within PhenoMeNal VRE. There is little effort



required to maintain a specific version, but if tools are further developed then the tests and containerisation needs to be kept in sync.

3. Development pipeline
 - a. Containerisation
 - b. Make available to services (like Galaxy, Luigi, Jupyter etc)
 - c. Test
 - d. CI (Jenkins) < could we move (a stripped-down version of) this to <https://travis-ci.org/> (like KubeNow has)

After M36:

A key objective from a technical sustainability perspective is to maintain an operational portal/gateway on a cloud provider, so that new users can be exposed and test launching VREs. Further, PhenoMeNal should preferably have a set of limited resources available so that new users apart from launching VREs also can test the features of the VRE without needing to purchase resources on a cloud provider. It would also be important to maintain functionality of portal and instantiation scripts (KubeNow and KubeClient) to instantiate PhenoMeNal VREs on the supported cloud providers. Apart from this, maintaining the continuous integration and testing and act upon changes in the dockerised tools that break any tests is of key importance to sustain a complete and working system after M36.

3.5 Technical sustainability scenarios

Sustainability scenario's

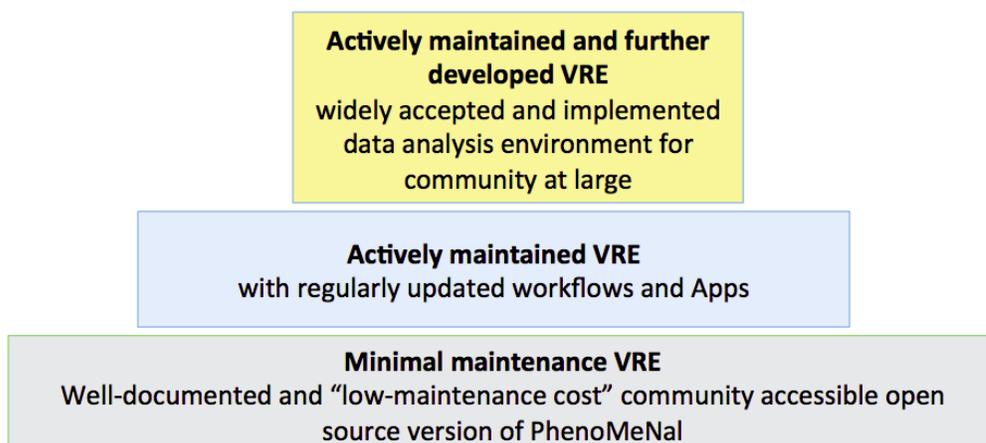




Figure 2. Technical sustainability scenarios

In D2.2, we presented different options/scenarios (VRE maintenance modes) depending on the degree of funding we are able to mobilise in the next years:

1) Minimal maintenance of PhenoMeNal VRE. In this scenario, active development of tools and workflows are not carried out, but the focus is on ensuring that the existing e-infrastructure components are available and can operate. An important task is the maintenance of the continuous integration server, and supervision of all existing projects ensuring that they conform to PhenoMeNal VRE standards and interoperability best practices over time as individual tools are further developed by external partners. Should there be major updates that are not possible to adapt for PhenoMeNal VRE, such tools would have to be removed from PhenoMeNal VRE. There would only be very limited public compute resources contributed by the PhenoMeNal project to the community, and users would have to rely on public cloud providers, internal datacentres with IaaS capabilities, or international consortia such as ELIXIR/EGI.

- Maintenance of e-infra to sustain PhenoMeNal over time and to assist users: indicative effort 2 FTE (200 k€/year)

2) Actively maintained VRE. This constitutes a scenario where the PhenoMeNal VRE is supported by an active community, that apart from maintaining the continuous integration server in scenario 1 above, also would provide regularly updated workflows and Apps and with active interaction and participation in tool development. The documentation is continuously improved, and outreach activities assist the community with further adopting the e-infrastructure.

- Maintenance of e-infra to sustain PhenoMeNal over time: 2 FTE
- Tool maintenance: 2 FTE
- Workflow maintenance: 1 FTE
- Documentation and outreach: 1 FTE
- Indicative effort (total): 6 FTE (600 k€/year)

3) Actively maintained and further developed VRE. This scenario includes, apart from all in scenarios 1 and 2 above, also future development of tools, new workflows in other scenarios, continuous work on data formats, refinement of virtual infrastructure components such as contextualisation, more demonstrations on projects within sensitive installations, and interaction and harmonisation with related projects (ELIXIR, EOSC, BBMRI and emerging projects on microservices in other domains e.g. genomics).



- Maintenance and further development of e-infrastructure: 5 FTE
- Tool maintenance and further development: 4 FTE
- Workflow maintenance and further development: 2 FTE
- Data format development: 2 FTE
- Additional demonstrators: 4 FTE
- Management, documentation, interaction and outreach: 3 FTE
- Indicative effort (total): 20 FTE (2 M€/year)

3.5.1 Funding sources for the scenarios

The core groups involved in the development of PhenoMeNal have been active to acquire (limited) national funding for the upkeep of PhenoMeNal according to the “minimal maintenance scenario” in the next 24 months after the end of the project. An overview of the national funding applications/grants is listed in Appendix 2.

To determine funding strategies for the more elaborate scenarios “Actively maintained VRE” and “Actively maintained and further developed VRE” we will organise a workshop in early autumn, when the H2020 Infrastructure programme 2018/2019 has been finalised and published. Ideally this would be a similar project in size as PhenoMeNal, now with a broader scope including the integration of proteomics, genomics and other omics data sources, allowing us to carry out the “Actively maintained and further developed VRE” scenario. However, we will also consider coordination and support actions, with a more limited budget, to carry out the “Actively maintained VRE” scenario.

Below is an update of current (international) initiatives we have undertaken to sustain PhenoMeNal, also underpinning the fact that for the sustainability of PhenoMeNal we are not betting on one horse.



4 METABOLOMICS ELIXIR USE CASE

A workshop in preparation for establishing a Metabolomics use case in ELIXIR was organised in collaboration with ELIXIR-NL¹ and ELIXIR-DE² on 25th April in Frankfurt, Germany. The one-day workshop showcased presentations on the metabolomics resources and services available through the ELIXIR nodes, ELIXIR use cases and their alignment with the ELIXIR platforms. This was followed by discussion sessions on identifying the needs and challenges of the metabolomics community, formulating use cases and their alignment with the existing ELIXIR platforms³. The highly interactive workshop resulted in a position paper for establishing a Metabolomics Use Case in ELIXIR. The latter was presented at the ELIXIR Heads of Nodes (HoN) meeting on June 13th 2017. The Use case was well received and people recognised the bottom up effort from the Metabolomics community. As a follow up, an application for the formal proposal for metabolomics as a Use Case for ELIXIR is currently under preparation. We have been invited to present this proposal at the next HoN meeting in September 2017.

5 IMPLEMENTATION NETWORK GO-FAIR/EUROPEAN OPEN SCIENCE CLOUD

Workshop on "Establishing a node for FAIRifying metabolomics and phenomics data in the European Open Science Cloud"

The European Open Science Cloud (EOSC) was launched by the European Commission (EC) in 2016 with the aim to ensure that science, business, and public services reap benefits of the big data revolution. Following the recommendations from our SAB about collaborating closely with the EOSC, a 2-day strategic workshop on "Establishing a node for FAIRifying metabolomics and phenomics data in the European Open Science Cloud" was hosted at the LU from 9th – 10th March 2017. The workshop was jointly hosted with the GO FAIR initiative. As an outcome of the workshop, on 10 March 2017, the 'Metabolomics Implementation Network' of the European Open Science Cloud was launched. The network aims to implement the FAIR data principles in the field of metabolomics in Europe. This will help scientists to exchange research data and workflows and to combine metabolomics data with other phenotypic data,

¹ <https://www.elixir-europe.org/about-us/who-we-are/nodes/netherlands>

² <https://www.elixir-europe.org/about-us/who-we-are/nodes/germany>

³ <https://www.elixir-europe.org/platforms>



enhancing the efficiency and impact of European life sciences research. To quote Professor Karel Luyben, chairman of the GO FAIR taskforce, “The GO FAIR initiative welcomes the Metabolomics Implementation Network as one of the first GO FAIR Implementation Networks. We actively support this network as a key element in the European Open Science Cloud.” The Metabolomics Implementation Network consists of key players in the European metabolomics community, including EMBL-EBI, Imperial College London, Leiden University, Oxford University, CIRMMMP, CNRS, University of Jena, Janssen Pharmaceuticals, DSM, the Hyve, and linked Data experts. The group signed the ‘Leiden Declaration on FAIR metabolomics’ and has set itself the following objectives:

- To form and develop the open Metabolomics Implementation Network on metabolomics data and services; we invite others to join;
- To collectively implement standards compliant with FAIR principles in the wider research community and to actively communicate these;
- To work closely together with other communities on better capturing and understanding phenotypes enabling integrated approaches;
- To support the GO FAIR initiative with the widespread implementation of the FAIR principles in the EOSC, contributing to a global open internet of FAIR data and services.

The Ministers of Germany, The Netherlands and France are in the process of establishing a European Coordination Office for the GO-FAIR initiative in Leiden, The Netherlands. One of the tasks of this office will be to interact with the different Implementation Networks. A long-term perspective is that these networks can be sustainably funded by so-called “cloud coins” that researchers will be awarded as part of their research grants, and need to be spend for FAIRyfyng their data and data stewardship.

Second stage proposal MetaStar for a metabolomics starting community on standardisation of wet analytical procedures building on PhenoMeNal in the H2020-INFRAIA-2017-1-two-stage call



The main goal of this project is to create a better harmonised virtual Metabolomics research infrastructure in Europe, by uniting the leading metabolomics facilities in Europe and at the same time educate the next generation of young researchers. MetaStar would build on the outcomes of PhenoMeNal with a strong focus and standardisation of wet analytical procedures. The project had a budget request of over € 5 million, for a duration of 48 months and included over 25 partners, including all major instrument vendors and academic outreach partners in Estonia, the Czech Republic, Greece and Portugal. Unfortunately, at the end of August the consortium was informed at the end of August that the proposal will not be funded in this call. However, writing this proposal has aligned future directions of the Metabolomics community working together with all major instrument vendors and has led to new collaborations with groups in Southern and Eastern Europe.



7 WORK PLAN

- Establish a European Metabolomics Foundation before M30 as a “driving force” to steer follow-up activities of the PhenoMeNal project and the Metabolomics Implementation Network of the EOSC.
- Organise a workshop in Autumn 2017 to discuss potential follow-up project(s) and/or coordination actions in calls described in the Horizon2020 work programme 2018/2019.
- Organise a follow-up workshop of the Metabolomics Implementation Network of the EOSC in Winter 2017/2018 to discuss the progress.
- Execute the Dissemination and Outreach activities to potential users and champions as detailed in D2.3
- Execute the activities according to the business plan, D2.4

8 CONCLUSIONS

Important steps towards the sustainability of PhenoMeNal have been taken:

- Preparations to set up a foundation are well under way, to steer and oversee the engagement with users and developers after the end of the project and to coordinate future national and international funding applications, as well as the EOSC Metabolomics Implementation Network
- Partners have secured enough national funding to realise the “Minimal maintenance scenario” for two years after the project, after which a substantial user base and successful follow-up funding applications should be realised.
- An official application for a Metabolomics Use Case in the ELIXIR work programme 2019-2023 is prepared for the Heads of Node meeting of ELIXIR in September



9 Appendix 1 Current installations of PhenoMeNal

9.1 Current installations of PhenoMeNal within the consortium

9.1.1 Uppsala University

Uppsala University is the site where the development of PhenoMeNal and the implementation in the metabolomics facility have taken place simultaneously.

Uppsala University uses PhenoMeNal VREs that are deployed on the national SNIC Science Cloud (<https://cloud.snic.se/>), on Google Cloud Platform, and on local servers. The CARAMBA platform (<http://www.caramba.clinic/>) for medical metabolomics uses PhenoMeNal tools and VRE for the majority of the ongoing studies. The main focus of CARAMBA is to perform pre-clinical and clinical research, targeting endogenous and exogenous molecules, with the overall goal to improve human health care and disease prevention. A secondary aim is implementing state-of-the-art as well as novel mass spectrometry and bioinformatics-based methods for use in routine clinical applications. PhenoMeNal has hence come to be an important component in the mass-spectrometry data analysis performed in the platform. Of special importance is the ability to run the VRE on in-house servers on sensitive data, and the ability to run identical runs on private and public cloud providers.

9.1.2 The Netherlands / Leiden University

In Leiden, the development of PhenoMeNal takes places in the research group. So far, PhenoMeNal has been tested in the production environment, the BioMedical Metabolomics Facility (BMFL), but is not implemented in the routine workflows. In the next period (M24-M30) PhenoMeNal will be fully implemented in the BMFL, in particular for the quality control tools.

The BMFL we work according to ISO17025 guidelines and are capable of running thousands of samples per month on the fully validated platforms. A crucial part of the workflow is assessing the quality of the measurements with our in-house developed quality control pipeline. The first version of the pipeline was developed in 2008, and has been improved and extended over the years. Back in 2008 the whole pipeline ran on a standard PC/laptop, the current version now runs on a dedicated workstation with 16-cores and 128GB of memory. In the near future we expect that large studies, with 500+ samples, will need even more compute power than currently is available at the BMFL.

For this reasons we have adopted the [tool development and deployment](#) strategy of PhenoMeNal in Leiden. For the quality control pipeline this meant:



- a full rewrite of the code in a different programming language (Python), to make it more open and accessible for others, in comparison to the old MatLab code.
- split into several small tools (modules) to be able to wrap in containers and to make them accessible in Galaxy (workflow engine).
- write tests to check the functionality of the individual tools.
- setup the continuous integration for automated testing and packaging.

By adopting this strategy at the BMFL we:

- are preparing our tools and pipelines to be deployable and therefore scalable to on-premises cloud and/or commercial cloud providers like OpenStack, Google or Amazon.
- we will be able to provide a per project instance of the pipeline(s), instead of a (single) shared environment.
- reduce cost by buying compute resources on-demand, instead of running and maintaining expensive servers 24/7.

Until now we have tested the PhenoMeNal services in the sponsored cloud and on local workstations using the MiniKube installation as described here: <https://github.com/phnmnl/phenomenal-h2020/wiki/QuickStart-Installation-for-Local-PhenoMeNal-Workflow>. The first tool, the initial step of the quality control pipeline, has been added to the PhenoMeNal release, and we expect the remaining tools to be added in the upcoming release.

9.1.3 EMBL-EBI

EMBL-EBI runs the reference PhenoMeNal public instance available on <https://public.phenomenal-h2020.eu/>, which always hosts the most recent release version of PhenoMeNal. This runs on a 6 node Kubernetes cluster deployed on the EMBL-EBI EMBASSY Cloud (OpenStack), with access to shared file system provisioned as GlusterFS. This instance, together with its sister development version available at <https://publicdev.phenomenal-h2020.eu/>, is monitored 24/7 to ensure uptime. Users can register for free account at the PhenoMeNal VRE Portal, at <https://portal.phenomenal-h2020.eu/>.

9.1.4 Germany (de.NBI)

The German Network for Bioinformatics Infrastructure (de.NBI) is operating a national cloud infrastructure in six locations running OpenStack cloud services across germany. Currently, there are no recommendations on how to lift complex application stacks on this german cloud. In PheNoMeNal, we are regularly testing the deployment to these OpenStack installations, to make sure that PhenoMeNal and de.NBI are and remain compatible. Our approach to metabolomics application stack deployment was part of



the de.NBI Summer School on Cloud Computing for Bioinformatics 2017 25th June - 1st July 2017 in Giessen.



10 Appendix 2 National policy measures supporting the development of PhenoMeNal in the next years

In D2.3 several national initiatives have been described, which (if funded) will contribute towards the sustainability of PhenoMeNal:

10.1.1 UK Institute for Health and Biomedical Informatics Research

Imperial is a member of the Francis Crick Institute together with the MRC, Wellcome Trust, CRUK, UCL and KCL. As part of this Imperial College has submitted an expression of interest to become a substantive site of the UK Institute for Health and Biomedical Informatics Research. The focus of this application is on new knowledge generation from public health, population, clinical, cellular and linked omics datasets (including PhenoMeNal analysis pipelines and infrastructure), skills and training applied to relevant patient needs based on the wealth of data generated across the IC healthcare landscape. This could sustain the instance of PhenoMeNal in the UK

10.1.2 UKMedBio infrastructure. UKMedBio / Medical Research Council

PhenoMeNal is using UKMedBio (MRC funded £6M infrastructure investment) as a test case for bringing the compute to the data and also using this large infrastructure to conduct scalability and optimisation testing. Imperial College London is involved in this project.

10.1.3 The Netherlands: X-omics initiative

In the Netherlands X-omics initiative has been formed, combining the strengths in genomics, proteomics and metabolomics and data analysis. A consortium has put forward a €16 million funding proposal to the National Science Foundation in the framework of the Roadmap Large Scale Research Infrastructures. In total a budget of € 45 million in 2017 and € 45 million in 2019 will be available for the Life Sciences. Leiden University is metabolomics innovation lead, additional partners include ErasmusMC, LUMC and Radboudumc. If funded, this will contribute to the sustainability of PhenoMeNal as the data analysis will build upon PhenoMeNal.



10.1.4 Swedish National Bioinformatics Infrastructure Sweden, (<http://nbis.se/>)

NBIS is a distributed national research infrastructure supported by the Swedish Research Council (Vetenskapsrådet), Science for Life Laboratory, all major Swedish universities and the Knuth and Alice Wallenberg Foundation, providing state-of-the-art bioinformatics to the Swedish life science researchers community. NBIS is also the Swedish contact point to the European infrastructure for biological information ELIXIR. NBIS can support e-infrastructure project based in an internal selection process. Right now NBIS supports the development of e.g. Human Metabolic Atlas. Status: NBIS has supported PhenoMeNal since start by employing software engineers that are hired to PhenoMeNal. Discussions on future sustaining PhenoMeNal in Sweden are ongoing.

10.1.5 Swedish Tryggve (<https://wiki.neic.no/wiki/Tryggve>)

Tryggve is a three-year project to establish a Nordic platform for collaboration on sensitive data, funded by NeIC and the ELIXIR nodes in Denmark, Finland, Norway and Sweden. The approach in the project is to utilise and connect existing capacities and services at the Nordic countries. The work is conducted in close collaboration with the user communities and Nordic e-infrastructure providers. Status: PhenoMeNal has been brought forward as a use case in the upcoming Tryggve 2 project, and is currently being evaluated. This would anchor PhenoMeNal in the nordic domain.