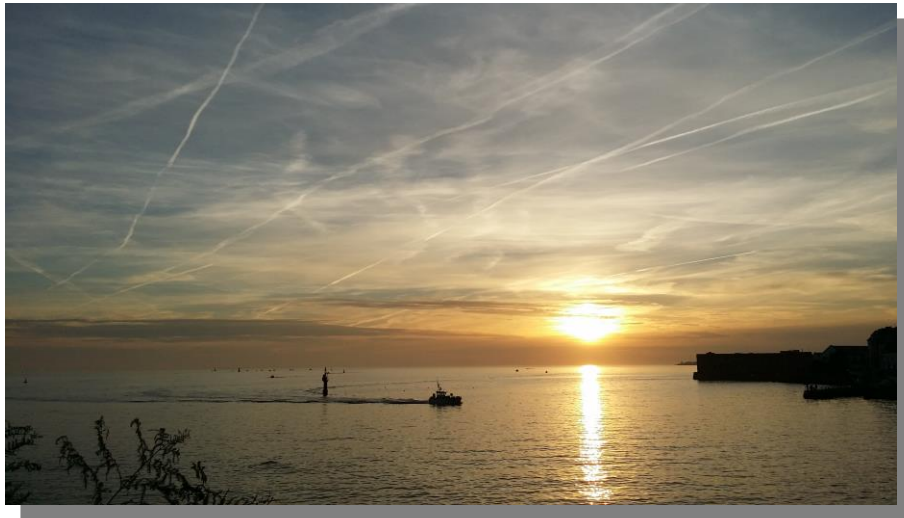
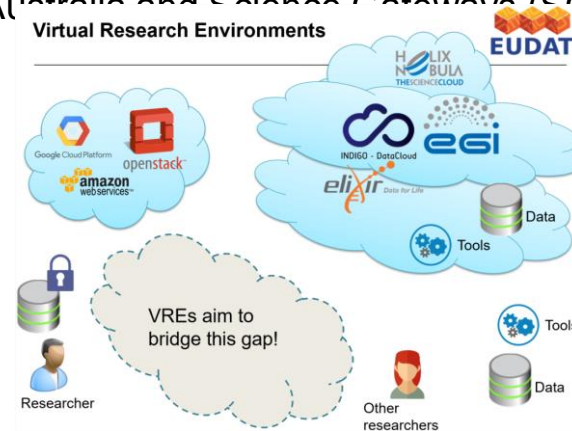


Petit tour d'horizon et comparatif des VREs



VRE Kesako

Virtual Research Environments (VREs) are increasingly being used to support a more dynamic approach to collaborative working in systematics and taxonomy. Researchers who are not co-located are seeking to work dynamically together at various scales from local to international. These shared infrastructures are funded as VREs in Europe, Virtual Laboratories (VLs) in Australia and Science Gateways (SGs) in the USA and all have similar objectives.



INDIANA UNIVERSITY

The Science Gateways Research Center researches, develops, and operates science gateways in collaboration with many clients and partners. Contact us for help building science gateways and deploying advanced software cyberinfrastructure for your community.

LEARN MORE

VRE Kesako

In fact VREs have a very long history in Europe. The UK Open University has used the term as an ICT support environment for students for many (at least 20) years. JISC (an organisation in UK supporting ICT across higher education) worked on VREs for higher education in 2004 and subsequently. Currently the EC (European Commission) favours the term VRE for systems or offerings that we also know as VLs or SGs. The term 'researcher workbench' has also been used widely since the 1980s and basically has the same concept as a VRE.

Key aspects seem to be (from the research I have done):

1. virtualisation (hiding complexity from the user);
2. access to useful resources such as datasets, software, computing power, instruments/detectors (the latter for control as well as data taking) and scholarly publications (including grey literature technical reports etc.) as well as collaboratively with other persons and organisations;
3. interoperability across resources;
4. support for the 'researcher workflow' from research idea (and checking the literature etc.) through observations/experiments to publication and subsequent discussion with citation and accreditation (maybe including management functions such as proposals and reporting to funders).
5. support for workflow composition (or even autonomic composition) of (2) and ideally deployment on virtualised resources (e.g. GRIDs, CLOUDs);

VRE Kesako

1. VRE = Virtual lab

- Online
- Central point of a distributed ecosystem, the e-Science one

2. e-Science

- Research + TIC
- « Research done through distributed global collaboration enabled by the internet, using very large data collections, terascale computing resources and high performance visualization » (Sir John Taylo – 2001)
- Three parts
 - Computing : HPC, grid, cloud
 - Storage
 - Collaboration

3. From infra to communities through softwares

- Research lifecycle

VRE Kesako

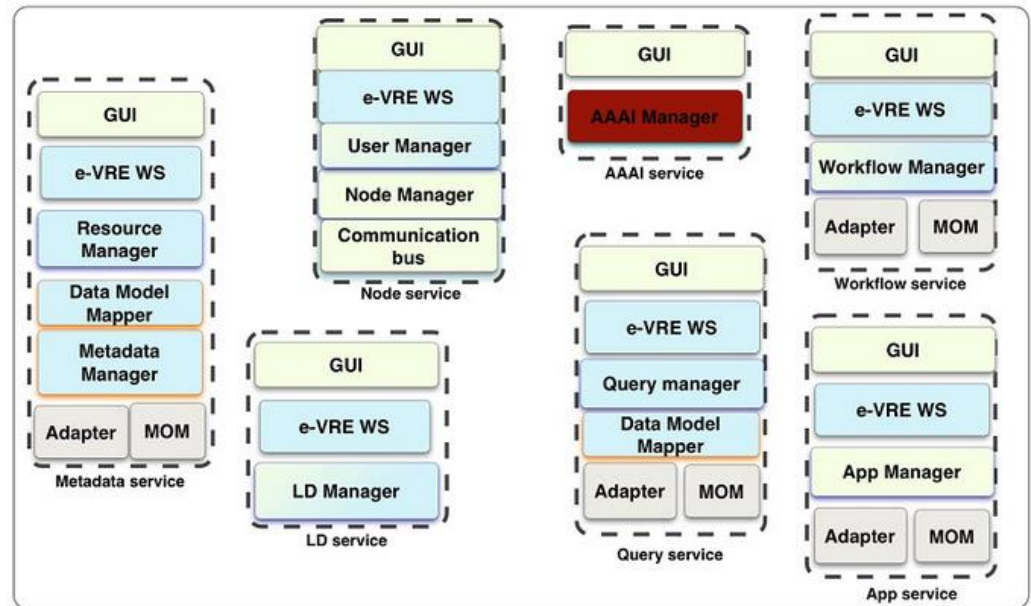
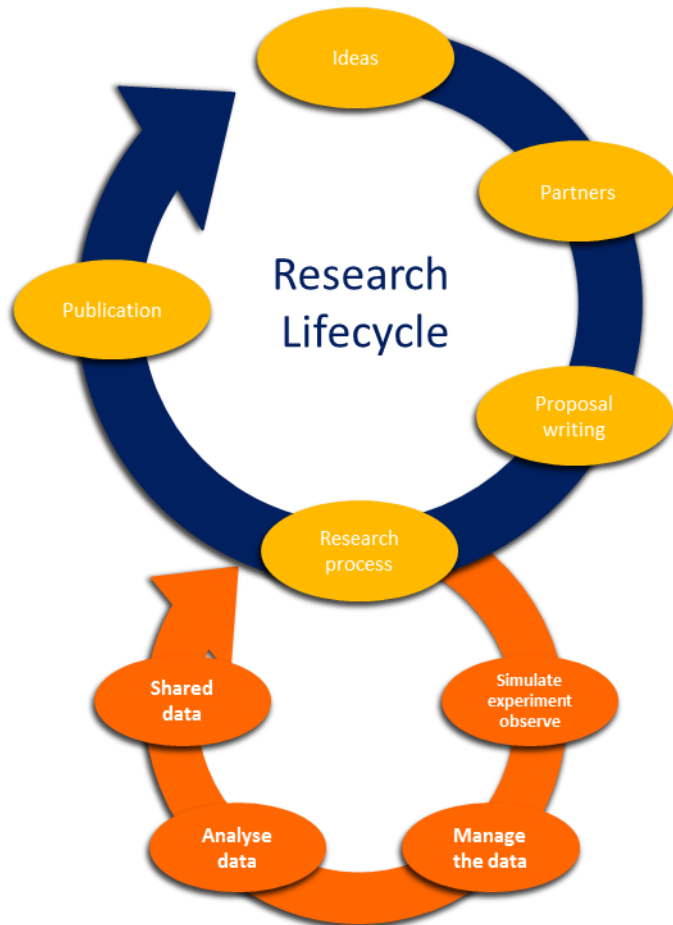
« Research done through distributed global collaboration enabled by the internet, using very large data collections, terascale computing resources and high performance visualization » (Sir John Taylo – 2001)

« A VRE comprises a set of online tools and other network resources and technologies interoperating with each other to facilitate or enhance the processes of research practitioners within and across institutional boundaries» [JISC definition](#)

« Virtual Research Environments are innovative, web-based, community-oriented, comprehensive, flexible, and secure working environments conceived to serve the needs of modern science» ([Candela et al. 2013](#))

- **Web** based
- Serve the **needs of researchers communities** (VRCs)
- Open and **flexible**

VRE, for which purpose ?



VRE « Tools »

1. Collaboration oriented tools

- Built on existing platform
 - HUBzero (using Joomla)
 - Sakai (a first demonstrator project so old ;))
 - [VRE4EIC](#) e-VRE
- Specific VRE framework
 - gCube
 - Microsoft Sharepoints VRE Toolkits
 - [OpenDreamkit](#) for mathematic
 - [Parthenos](#) for Humanities
 - [Phenomenal](#) on-demand VRE for metabolomics
 - [Apache Airavata](#) software framework
 - VRE OSF ?

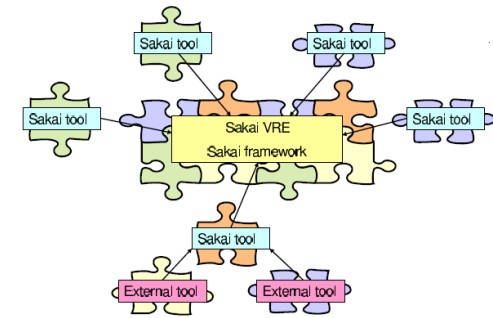
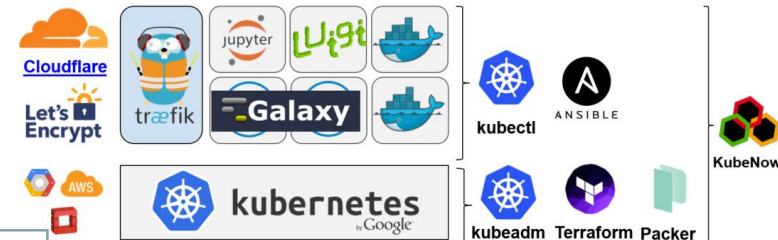


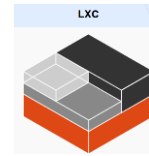
Figure 1. Pluggable Sakai VRE architecture.

Yang et al. 2006 (Integration of Existing Grid Tools in Sakai VRE)



Apache Airavata is a software framework which is dominantly used to build Web-based science gateways and assist to compose, manage, execute and monitor large scale applications and workflows on distributed computing resources such as local clusters, supercomputers, national grids, academic and commercial clouds. Airavata mainly supports long running applications and workflows on distributed computational resources.





Singularity



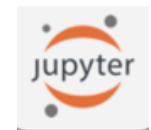
microservices



Dependencies management



Workflow manager



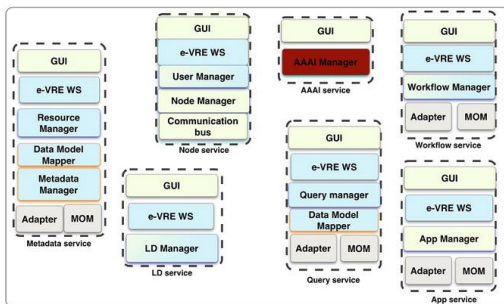
Analyze through GUI

VRE « Tools »

1. Collaboration oriented tools

- Built on existing platform
 - HUBzero (using Joomla)
 - Sakai (a first demonstrator project so old ;)
 - [VRE4EIC](#) e-VRE
- Specific VRE framework

- gCube
- Microsoft Sharepoints VRE Toolkits
- [OpenDreamkit](#) for mathematic
- [Parthenos](#) for Humanities
- Phenomenal on-demand VRE for metabolomics
- [Apache Airavata](#) software framework
- VRE OSF ?



ANSIBLE

SALTSTACK

Terraform

manage « infrastructure as code »

Cloud

infrastructure

VRE « Tools »

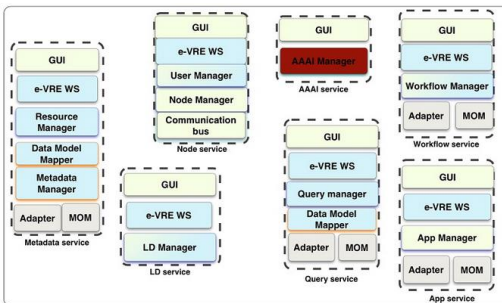
1. Collaboration oriented tools

- Built on existing platform



- HUBzero (using Joomla)
- Sakai (a first demonstrator project so old ;))
- [VRE4EIC](#) e-VRE

- Specific VRE framework



- gCube
- Microsoft Sharepoints VRE Toolkits
- [OpenDreamkit](#) for mathematic
- [Parthenos](#) for Humanities
- Phenomenal on-demand VRE for metabolomics
- [Apache Airavata](#) software framework
- VRE OSF ?

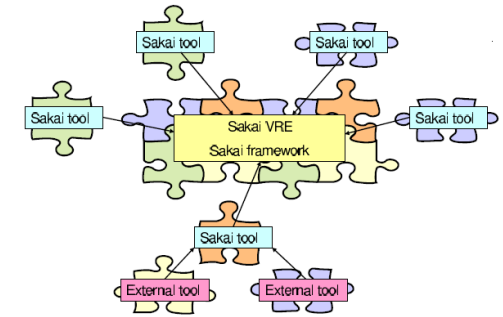
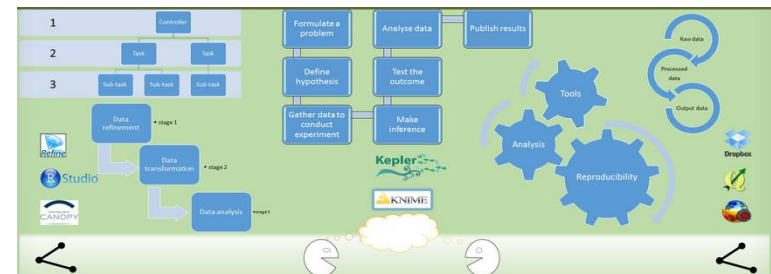


Figure 1. Pluggable Sakai VRE architecture.

Yang et al. 2006 (Integration of Existing Grid Tools in Sakai VRE)

2. VRE and Virtual / Learning Environments / Virtual Research Communities (VRCs)



VRE components

1. Stack

- Comparing gcube / HUBzero and [Phenomenal VRE](#)



2. Links

- Gateway -> VREs

3. Granulometry

- Size / subject of communities ?
- Genericity of tools as VRE components

4. Important questions

- Who will use it? Not only needs are important but also particularities of communities (ie Biologist don't code)

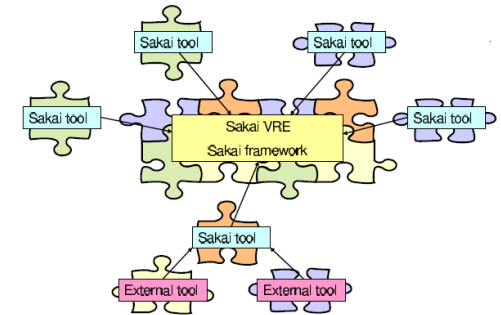
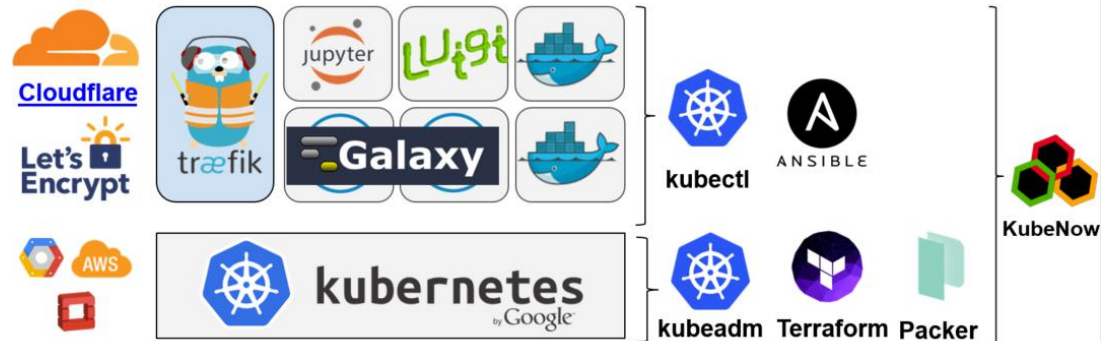
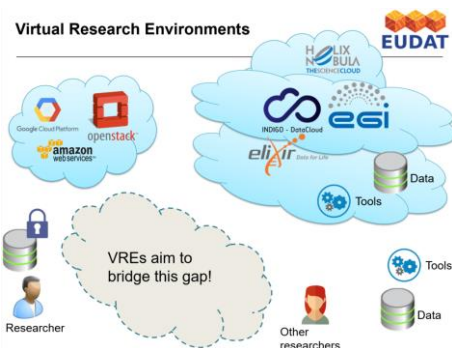


Figure 1. Pluggable Sakai VRE architecture.

Yang et al. 2006 (Integration of Existing Grid Tools in Sakai VRE)



VRE Kesako

Please, do take into account that there six projects officially supported by the EC in 2016 to develop VREs (<https://ec.europa.eu/programmes/horizon2020/en/news/six-new-projects-e-i...>):

- * BlueBRIDGE - Building Research environments for fostering Innovation, Decision making, Governance and Education to support Blue growth
 - * MuG - Multi-Scale Complex Genomics
 - * OpenDreamKit - Open Digital Research Environment Toolkit for the Advancement of Mathematics
 - * VI-SEEM - VRE for regional Interdisciplinary communities in Southeast Europe and the Eastern Mediterranean
 - * VRE4EIC - A Europe-wide Interoperable Virtual Research Environment to Empower Multidisciplinary Research Communities and Accelerate Innovation and Collaboration
 - * West-Life - World-wide E-infrastructure for structural biology
- * The *EVER-EST Horizon 2020* project will create a state of the art virtual research environment (VRE) focused on the earth sciences. <https://ever-est.eu/>



Wf4Ever Research Object Model 1.0

EVER-EST example

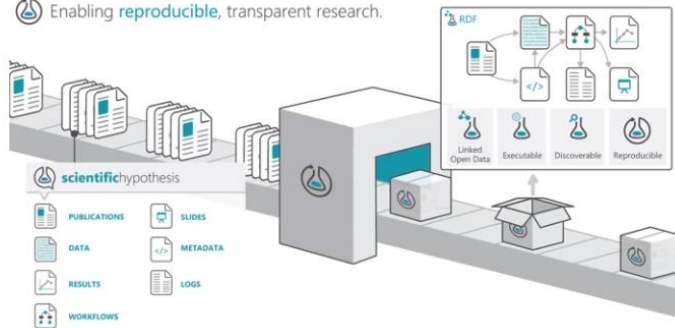
H2020 project for Earth-science

WHAT IS EVER-EST?

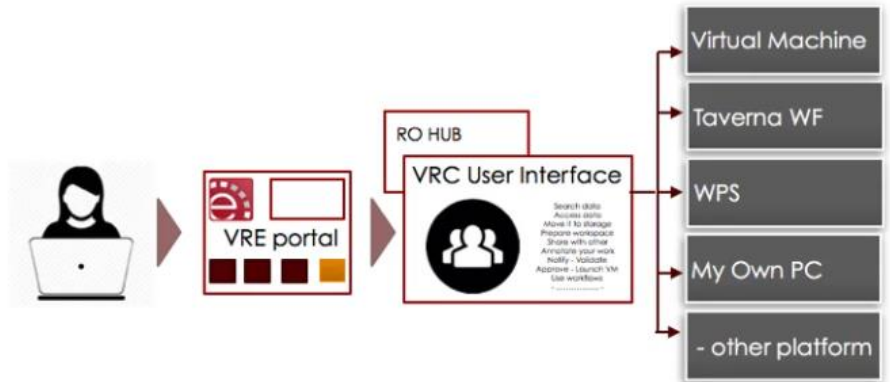


The solution is based on the integration of Earth Science e-infrastructures technologies - developed over 15 years - with the innovative concept of Research Object

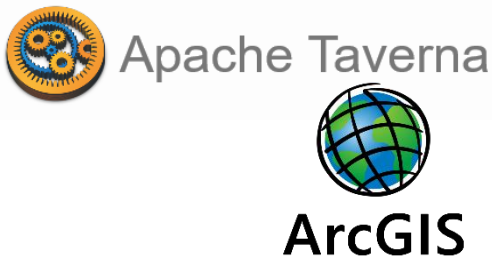
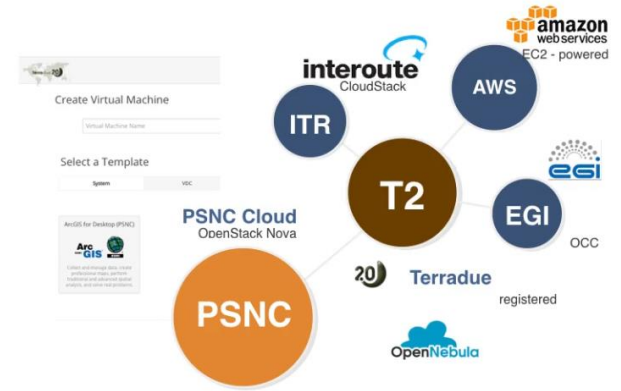
Enabling **reproducible**, transparent research.



The system can be easily integrated with different mechanisms to process Earth Science data



It takes full advantage of ES data cloud processing capabilities



Wf4Ever Research Object Model 1.0

EVER-EST example

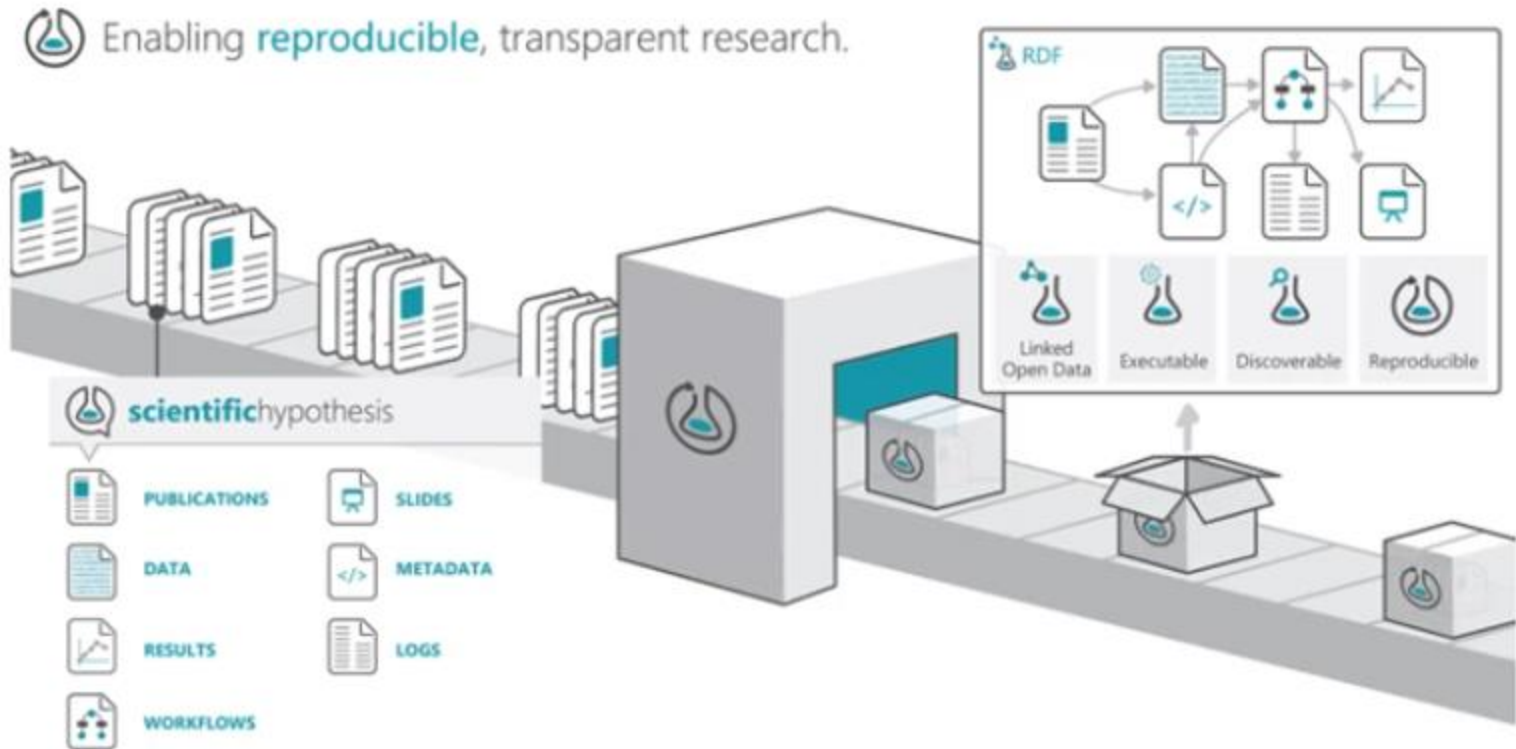
EVER-COL-040	Project Management Tools	Highly Desirable
Requirement	<i>EVER-EST e-infrastructure shall provide project planning and management tools to be used by the VRCs communities that includes presentation and editing of Gantt charts, reporting and recording of minutes and actions.</i>	
Source	VRE Use Cases (VM, SS)	
Verification Method	Verification Methods: Validated	
Note	not implemented yet, VRCs use as <i>project planning and management tools ALFRESCO</i>	

“**VRE as a Services**” is the response to the impelling need for an **e-Science** infrastructure supporting: **reproducibility, reuse and interdisciplinarity** insights and cross-fertilisation of scientific investigation and scholarly data, by developing and operating an Virtual Research Environment sustaining **Open Science** and implementing **FAIR Guiding Principles for scientific data management and stewardship**

Open Science covers the following aspects:

- Open Discovery, tools able to increase the efficiency of research as well as of its diffusion;
- Open Access to scientific inputs and outputs as research outcomes;
- Open Data, freely to available to everyone to use and publish, without restrictions from copyright, patents or other mechanisms of control;
- Open Methodology related to research strategy that outlines the way in which research is to be undertaken;
- Open Educational, allowing the broaden access to the learning and training traditionally offered through formal education systems.

EVER-EST example -> what is particularly of interest



RO must be associated to metadata -> here we see the interest of the « complete » of EML as this standard seems to be perfect to inform about each component of ROs

EVER-EST example -> what is particularly of interest



- R / RStudio
- Galaxy-E



RO must be associated to metadata -> here we see the interest of the « complete » of EML as this standard seems to be perfect to inform about each component of ROs

Some issues for me

Tool/Software/Script

Developped by communities
of practices

-free

-usable only in local PC

+

SaaS

PaaS

IaaS

Provided by private
companies

Tool/Software/Script

SaaS

PaaS

IaaS

Developped by international
project with « business »
delivrables such as H2020

-not free

-usable only through private
compagnies services

Some issues for me

Tool/Software/Script

Developped by communities of practices
-free
-usable only in local PC

Gcube packaging through SAI

Hidden source code

Gcube packaged tool

SaaS

PaaS

IaaS

+



SaaS

PaaS

IaaS

Provided by D4Science



EUBrasil OpenBio



Developped by international project using D4science
-free ?
-possibility to create an account ?
-usable only through D4science VREs

Statistical Algorithms Importer (SAI)



Ref: OpenScience Data Analytics Technologies, AGINFRA+ wiki

https://support.d4science.org/projects/aginfraplus_wiki/wiki/D31_-_Open_Science_Data_Analytics_Technologies#SAI

EU projects related to gcube:

- IMARINE - Data e-Infrastructure Initiative for Fisheries Management and Conservation of Marine Living Resources (283644)
- BlueBRIDGE - Building Research environments for fostering Innovation, Decision making, Governance and Education to support Blue growth (675680)
- EGI-Engage - Engaging the EGI Community towards an Open Science Commons (654142)
- PARTHENOS - Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies (654119)
- D4SCIENCE-II - Data Infrastructure Ecosystem for Science (239019)
- EUBRAZILOPENBIO - EU-Brazil Open Data and Cloud Computing e-Infrastructure for Biodiversity (288754)
- ENVRI PLUS - Environmental Research Infrastructures Providing Shared Solutions for Science and Society (654182)
- AGINFRA PLUS - Accelerating user-driven e-infrastructure innovation in Food Agriculture (731001)
- D4SCIENCE - Distributed colLaboratories Infrastructure on Grid ENabled Technology 4 Science (212488)
- ENVRI - Common Operations of Environmental Research Infrastructures (283465)
- SoBigData - SoBigData Research Infrastructure (654024)

DataMiner

“Required packages are assumed to be preinstalled on the backend system”

Java / Maven / Eclipse -> not for ecologists who develop scripts, need for IT JAVA guys



Ref: OpenScience Data Analytics Technologies, AGINFRA+ wiki
[https://support.d4science.org/projects/aginfraplus_wiki/wiki/D31 - Open Science Data Analytics Technologies#SAI](https://support.d4science.org/projects/aginfraplus_wiki/wiki/D31_-_Open_Science_Data_Analytics_Technologies#SAI)

EU projects related to gcube:

- IMARINE - Data e-Infrastructure Initiative for Fisheries Management and Conservation of Marine Living Resources (283644)
- BlueBRIDGE - Building Research environments for fostering Innovation, Decision making, Governance and Education to support Blue growth (675680)
- EGI-Engage - Engaging the EGI Community towards an Open Science Commons (654142)
- PARTHENOS - Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies (654119)
- D4SCIENCE-II - Data Infrastructure Ecosystem for Science (239019)
- EUBRAZILOPENBIO - EU-Brazil Open Data and Cloud Computing e-Infrastructure for Biodiversity (288754)
- ENVRI PLUS - Environmental Research Infrastructures Providing Shared Solutions for Science and Society (654182)
- AGINFRA PLUS - Accelerating user-driven e-infrastructure innovation in Food Agriculture (731001)
- D4SCIENCE - Distributed colLaboratories Infrastructure on Grid ENabled Technology 4 Science (212488)
- ENVRI - Common Operations of Environmental Research Infrastructures (283465)
- SoBigData - SoBigData Research Infrastructure (654024)



“Required packages are assumed to be preinstalled on the backend system”

Java / Maven / Eclipse -> not for ecologists who develop scripts, need for IT JAVA guys

D4Science Domains and Gateways

D4Science serves different domains

Domain	Percentage
Marine	60.4%
Cultural Heritage	21.3%
Social/Mining	8.1%
ICT	8.1%
Environmental	0%
Digital Library	0%
Agriculture	0%

Gateways shown: AGINFRA Gateway, ARIADNE Gateway plus, BlueBRIDGE Gateway, SoBigData Gateway, SOLID EARTH Gateway, OpenAIRE CONNECT Gateway, PARTHENOS Gateway, PerformFISH Gateway, GEMex, marine.

<https://services.d4science.org/thematic-gateways>

AGINFRA Realising a Science Gateway for the Agri-food: the AGINFRA PLUS Experience - M. Assante WWW.PLUS.AGINFRA.EU

Statistical Algorithms Importer (SAI)



Ref: OpenScience Data Analytics Technologies, AGINFRA+ wiki

https://support.d4science.org/projects/aginfraplus_wiki/wiki/D31_-_Open_Science_Data_Analytics_Technologies#SAI

DataMiner



“Required packages are assumed to be preinstalled on the backend system”

Java / Maven / Eclipse -> not for ecologists who develop scripts, need for IT JAVA guys

Statistical Algorithms Importer (SAI)

“an” ~Rstudio like “interface that allows scientists to easily and quickly import R scripts onto DataMiner”

accessible via the WPS standard

“The algorithms installer manage a list of users with visibility rights on an algorithm”



Ref: OpenScience Data Analytics Technologies, AGINFRA+ wiki
https://support.d4science.org/projects/aginfraplus_wiki/wiki/D31 - Open Science Data Analytics Technologies#SAI



“Required packages are assumed to be preinstalled on the backend system”

Java / Maven / Eclipse -> not for ecologists who develop scripts, need for IT JAVA guys

“an” ~Rstudio like “interface that allows scientists to easily and quickly import R scripts onto DataMiner”
 accessible via the WPS standard

“The algorithms installer manage a list of users with visibility rights on an algorithm”



Workflow orientation -> Limitation for HPC or not open source

Statistical Algorithms Importer (SAI)



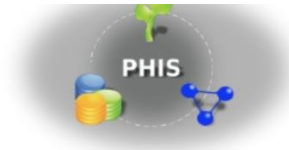
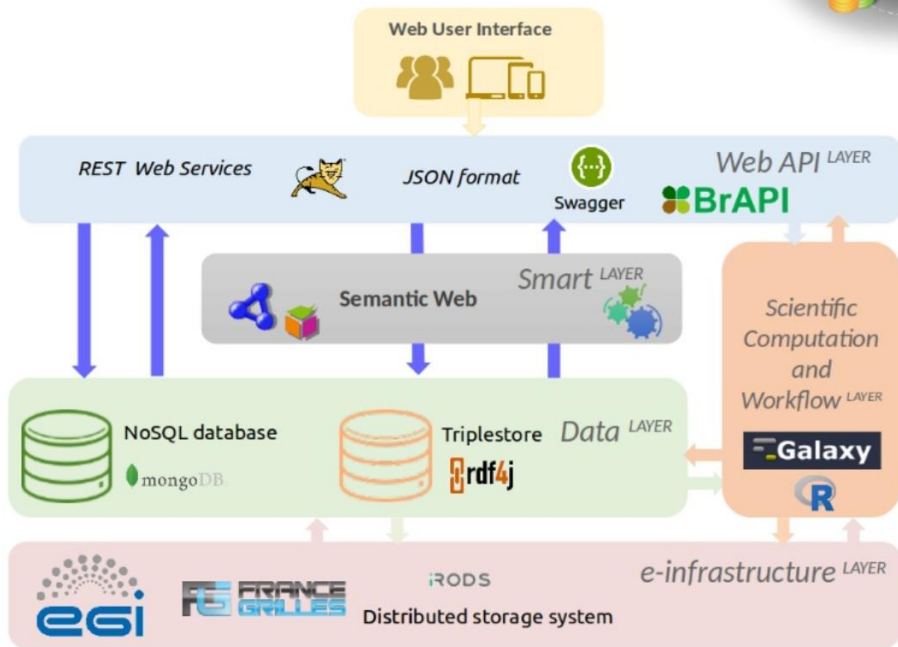
Ref: OpenScience Data Analytics Technologies, AGINFRA+ wiki

https://support.d4science.org/projects/aginfraplus_wiki/wiki/D31 - Open Science Data Analytics Technologies#SAI

And now, linking Galaxy to D4Science through RapidMiner WPS services available on Galaxy and more!

OpenSILEX - PHIS

Architecture



The Food Security VRE

What are the functionalities ?

Food Security VRE	
Data Access - Shared Workspace - Catalogue	Data Visualization - Visualization tool
Data Analytics - Rstudio - Jupyter Lab - Galaxy - Dataminer	Semantics - Vocbench - Yam++ - Silk

EGI services



<http://www.plus.aginfra.eu/>
<https://aginfra.d4science.org/>



<http://www.opensilex.org/>
<https://github.com/OpenSILEX>
<http://phis.inra.fr/>

<https://www.slideshare.net/aginfra/data-intensive-agricultural-sciences-requirements-based-on-aginfra-project-and-high-throughput-phenotyping-infrastructure>

Vincent Negre **Data intensive agricultural sciences : requirements based on Agintra+ Project and high throughput phenotyping infrastructure**

Statistical Algorithms Importer (SAI)

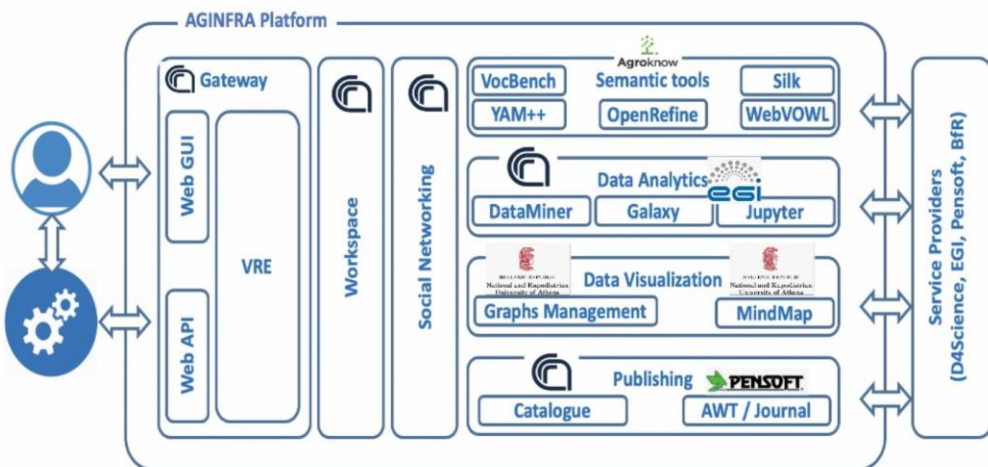


Ref: OpenScience Data Analytics Technologies, AGINFRA+ wiki

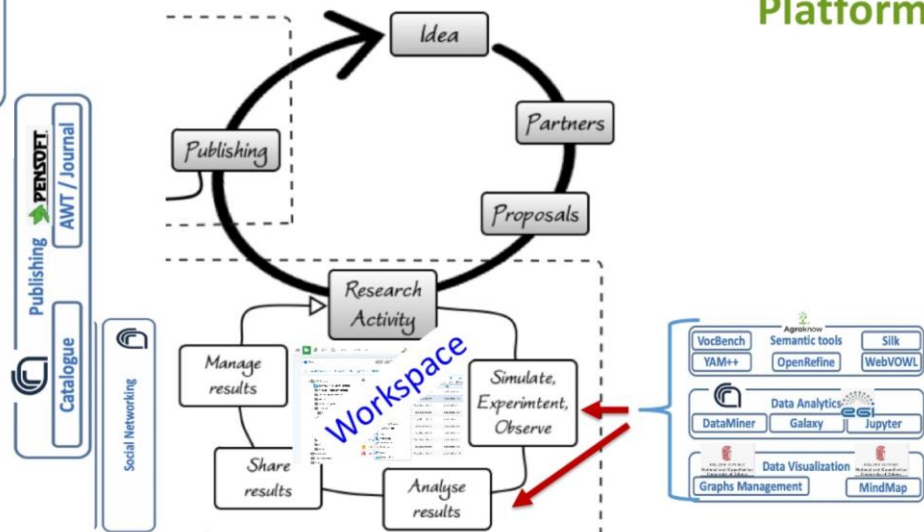
https://support.d4science.org/projects/aginfraplus_wiki/wiki/D31_-_Open_Science_Data_Analytics_Technologies#SAI

And now, linking Galaxy to D4Science through RapidMiner WPS services available on Galaxy and more!

AGINFRA Plus Platform Architecture



Research Life Cycle VS AGINFRA Platform



Realising a Science Gateway for the Agri-food: the AGINFRA PLUS Experience - M. Assante

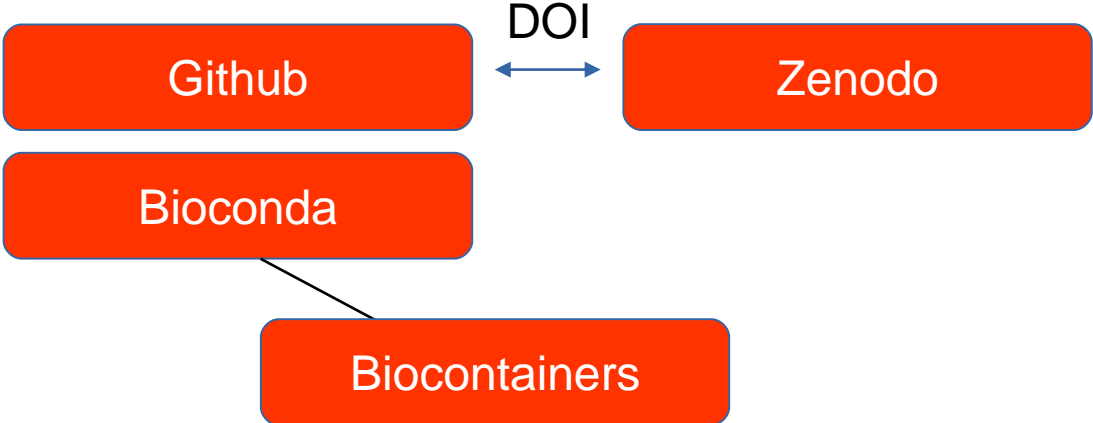
WWW.PLUS.AGINFRA.EU

https://www.slideshare.net/aginfra/realising-a-science-gateway-for-the-agrifood-the-aginfraplus-experience?next_slideshow=1

Massimiliano Assante CNR IWGS 2019 Realising a Science Gateway for the Agri-food: the AGINFRAplus Experience

One personal proposal for ecology communities

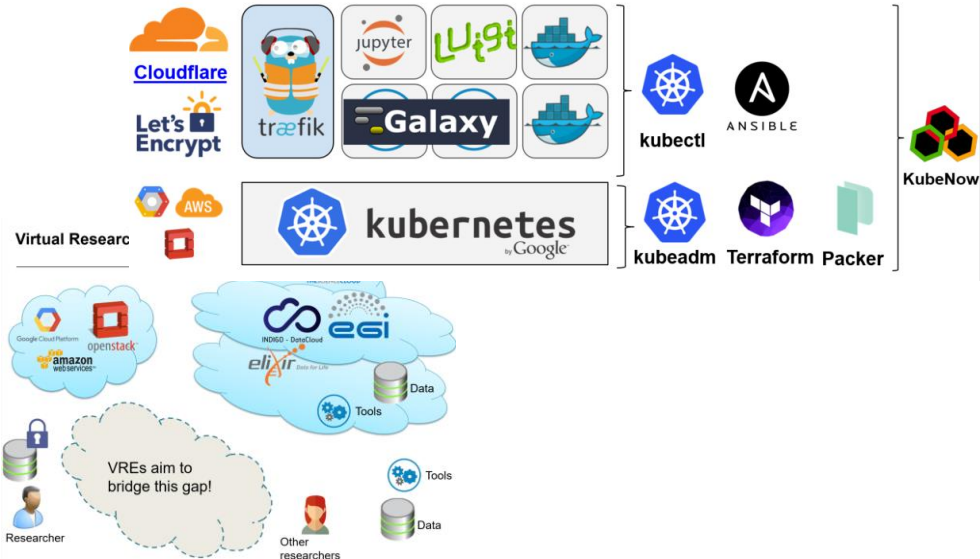
Tool/Software/Script



Other initiatives to consider :

- software heritage
- Easybuild
- Linux Guix

One free cloud based GUI solution : Galaxy



Microservices



Choose the services you want and price

- local (linux/windows/...)
- server, HPC or cloud
- command line or GUI



User Login

[What is GenomeSpace?](#) [Tools](#) [Recipes](#) [Documentation](#) [Developers](#) [Support](#) [About](#)



The GenomeSpace project is ending

The GenomeSpace project ends on November 15, 2019 due to expiration of its NHGRI funding and we will be shutting down the GenomeSpace servers on that date. We would like to thank all GenomeSpace users for their support and for all the important science they have done on the platform over the last nine years.

[GenomeSpace-enabled tools](#) are separately funded and will continue to be available subject to their own lifecycles, but any connection to GenomeSpace or GenomeSpace-hosted files will cease with the termination of the servers. If you do not remember your login credentials for any individual tools that you rely on, you should use the 'forgot password' links on their websites to reset your login credentials.

NOTE: We are unable to archive any data you have in your GenomeSpace account. Therefore, if you wish to retain copies of such data **you should transfer them to your own storage** from GenomeSpace before the termination date. After that time they will be deleted from the GenomeSpace servers as part of the shutdown process. Data that you host on other platforms (Dropbox, S3, Google Drive) and have connected to your GenomeSpace account will not be affected.

If you have any questions, please contact the GenomeSpace team at gs-help@broadinstitute.org.

The GenomeSpace Team

Merci de votre attention

