



***AGINFRA+: Virtual Research
Environments to Support Agriculture
and Food Research Communities***

L. Candela, P. Orviz

 eosc-hub.eu

Dissemination level: Public

 [@EOSC_eu](https://twitter.com/EOSC_eu)

- **Aim: Deploy a DataMiner cluster and make it available to D4Science communities**
- **Maturity (TRL) level: TRL9**
 - EGI FedCloud
- **EOSC services:**
 - EOSC Federated Authentication
 - EOSC Accounting Service (through EGI FedCloud)
 - EOSC Monitoring Service

- **Cloud Compute service**
 - *IFCA-LCG2 cloud site since April 17th*
 - *SLA (01/04 -> 31/12)*
 - <https://documents.egi.eu/public/ShowDocument?docid=2875>
- **EOSC Accounting Service**
 - *Sum Elapsed * Num Processors: 67980 hours*
 - May: 36 tasks
 - 15K hours, 241Gbytes
 - June: 25 tasks
 - 25K hours, 170Gbytes
 - July: 1 task
 - 14K hours, 65Gbytes

- **EOSC Monitoring Service**
 - Availability/Uptime
 - **JIRA issue:** <https://jira.eosc-hub.eu/browse/EOSCWP10-117>
 - Specific Nagios probes will be implemented to check the performance of the DM cluster
 - Currently on discussion if this is feasible through DM cluster API
 - **JIRA issue:** <https://jira.eosc-hub.eu/browse/EOSCWP10-116>

Summary - Roadmap status

<p><i>Quarter 1</i></p>	<ul style="list-style-type: none"> • Identify the initial resource capacity [DONE] • Engage with EOSC-hub providers [DONE] • Deploy DM cluster in a single cloud provider [IN PROGRESS][DONE] • Enabling EOSC-hub monitoring and accounting [DONE]
<p><i>Quarter 2</i></p>	<ul style="list-style-type: none"> • Quarterly review of performance thresholds <ul style="list-style-type: none"> ◦ Analytics tasks increased 10% [May: 36, June: 3, July: 1] ◦ Availability (uptime) > 95% [DONE] • Availability of monitoring probes for checking the status of the DM cluster (through API) [IN PROGRESS, JIRA's EOSCWP10-117]
<p><i>Quarter 3</i></p>	<ul style="list-style-type: none"> • Quarterly review of performance thresholds (see Q2) • Integration of monitoring probes in EOSC Monitoring service
<p><i>Quarter 4</i></p>	<ul style="list-style-type: none"> • Quarterly review of performance thresholds (see Q2) • Review of EAP experience and assessment of operational continuity of the application • Registration of AGINFRA+ VREs in the EOSC Portal






Community Requirements DB

[AGINFRA+ entry](#) created:

- Capacity Reqs
- Technical Reqs: JIRA issues
 - *EAP services not listed in JIRA, how can they be tracked?*

EOSC Portal

Not yet there, *should this be done by the EAP (through provider portal)?*

Requirement number	Requirement title	Link to Requirement JIRA ticket	Source Use Case
Example	EOSC-hub to provide an FTS data transfer service	 EOSCWP10-21 - EOSC-hub to provide an FTS data transfer service IN PROGRESS	UC1
RQ1	DM cluster deployment in the EGI Cloud Compute service	 DONE	UC1
RQ2	Tracking compute consumption		UC2
RQ3	Get performance metrics: number of analytics tasks	 n° analytics tasks May: 36 June: 3	UC3
RQ4	Get performance metrics: availability / uptime (percentage)	 EOSCWP10-117 - Get performance metrics: availability / uptime for AGINFRA+ service TODO	UC3
RQ5	Display of performance metrics in Nagios	 EOSCWP10-116 - Display of performance metrics of DM cluster in EOSC monitoring service TODO	UC3

Thank you for your attention!

Questions?



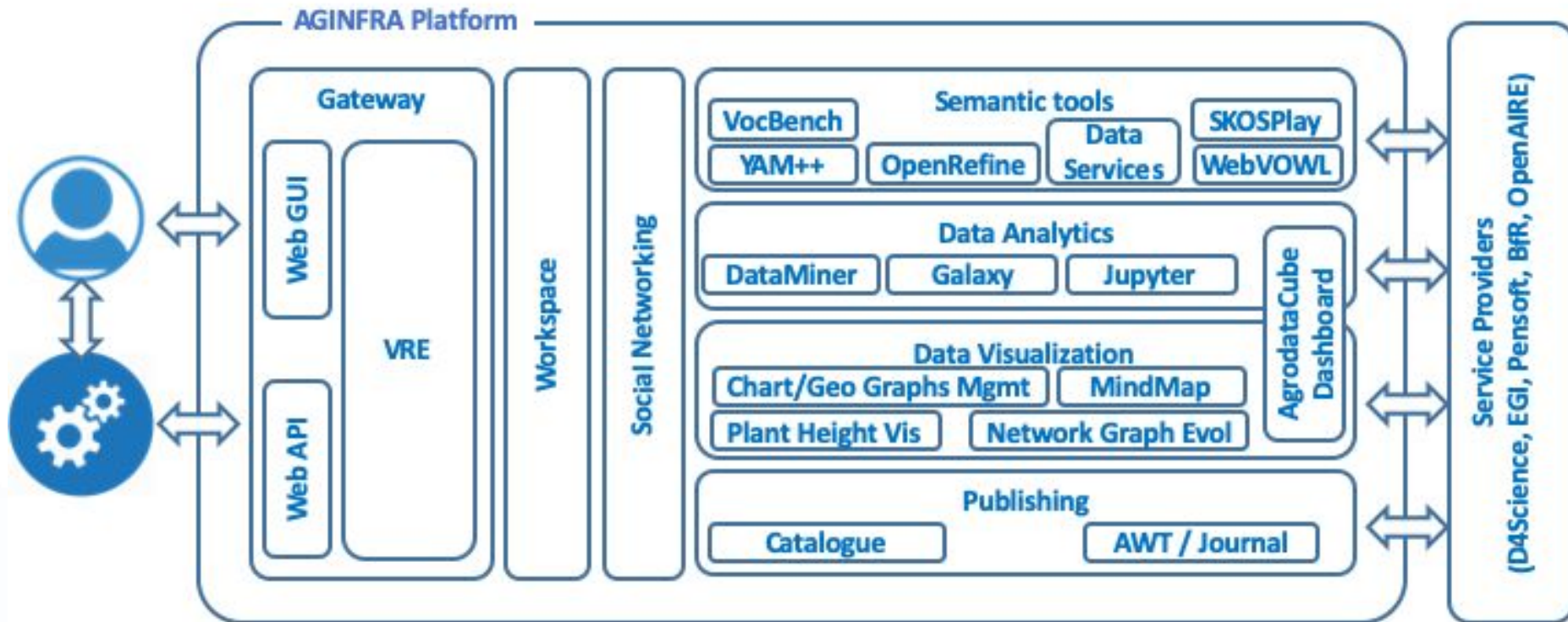
EOOSC-hub

 eosc-hub.eu  [@EOOSC_eu](https://twitter.com/EOSC_eu)

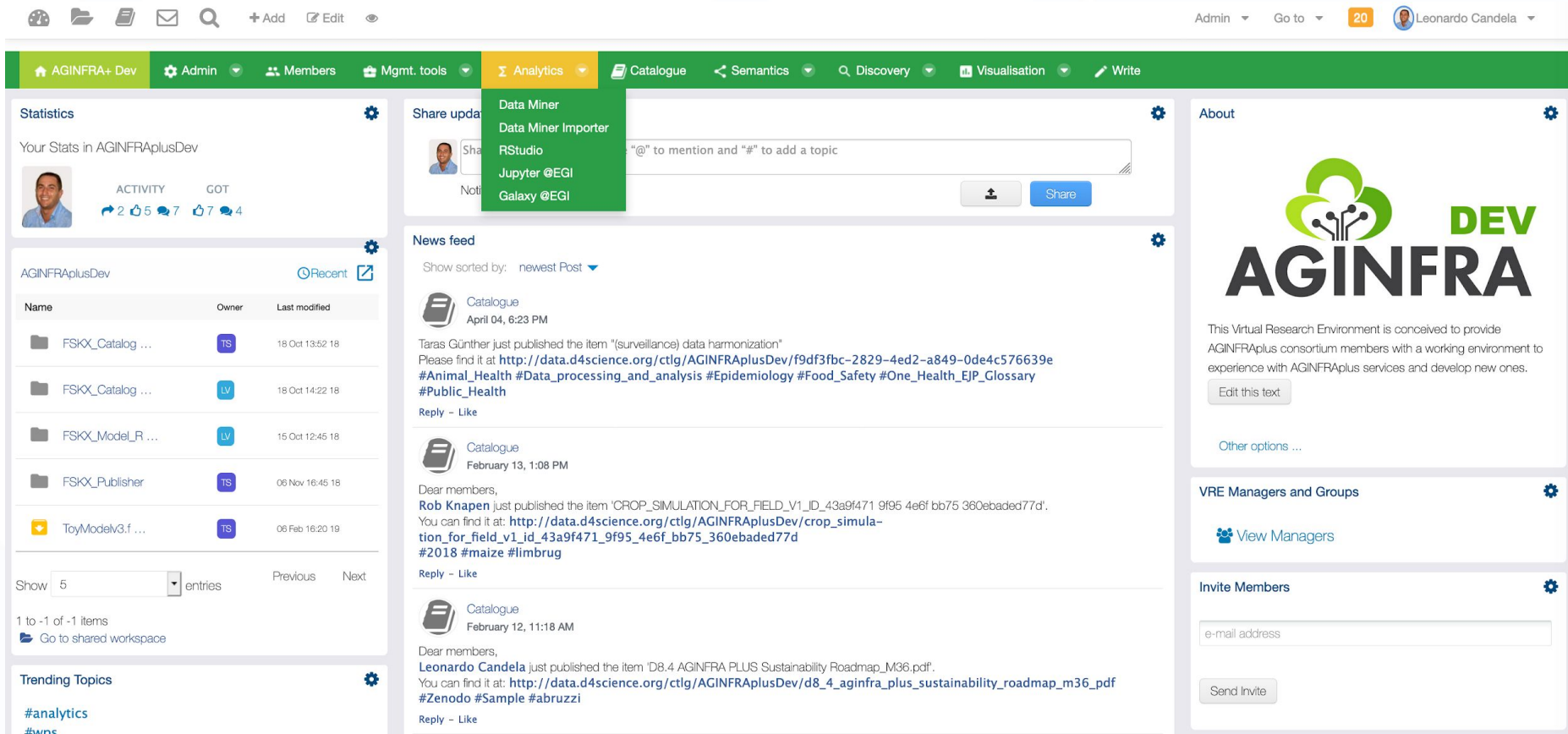


This material by Parties of the EOOSC-hub Consortium is licensed under a Creative Commons Attribution 4.0 International License.

<https://aginfra.d4science.org/home>



<https://aginfra.d4science.org/home>



The screenshot shows the AGINFRA+ Dev dashboard interface. At the top, there is a navigation bar with icons for home, files, documents, mail, search, and user actions (+ Add, Edit, eye). The main navigation bar is green and contains: AGINFRA+ Dev, Admin, Members, Mgmt. tools, Analytics (highlighted), Catalogue, Semantics, Discovery, Visualisation, and Write.

Statistics: Your Stats in AGINFRAplusDev. Includes a profile picture and activity metrics: ACTIVITY (2 likes, 5 comments, 7 shares) and GOT (7 likes, 4 comments).

AGINFRAplusDev Table:

Name	Owner	Last modified
FSKX_Catalog ...	TS	18 Oct 13:52 18
FSKX_Catalog ...	LV	18 Oct 14:22 18
FSKX_Model_R ...	LV	15 Oct 12:45 18
FSKX_Publisher	TS	06 Nov 16:45 18
ToyModelv3.f ...	TS	06 Feb 16:20 19

Showing 5 entries. Previous Next. 1 to -1 of -1 items. Go to shared workspace.

Trending Topics: #analytics, #wps

Share update: A dropdown menu is open showing options: Data Miner, Data Miner Importer, RStudio, Jupyter @EGI, and Galaxy @EGI. The input field contains "@@ to mention and #" to add a topic. Buttons for upload and Share are visible.

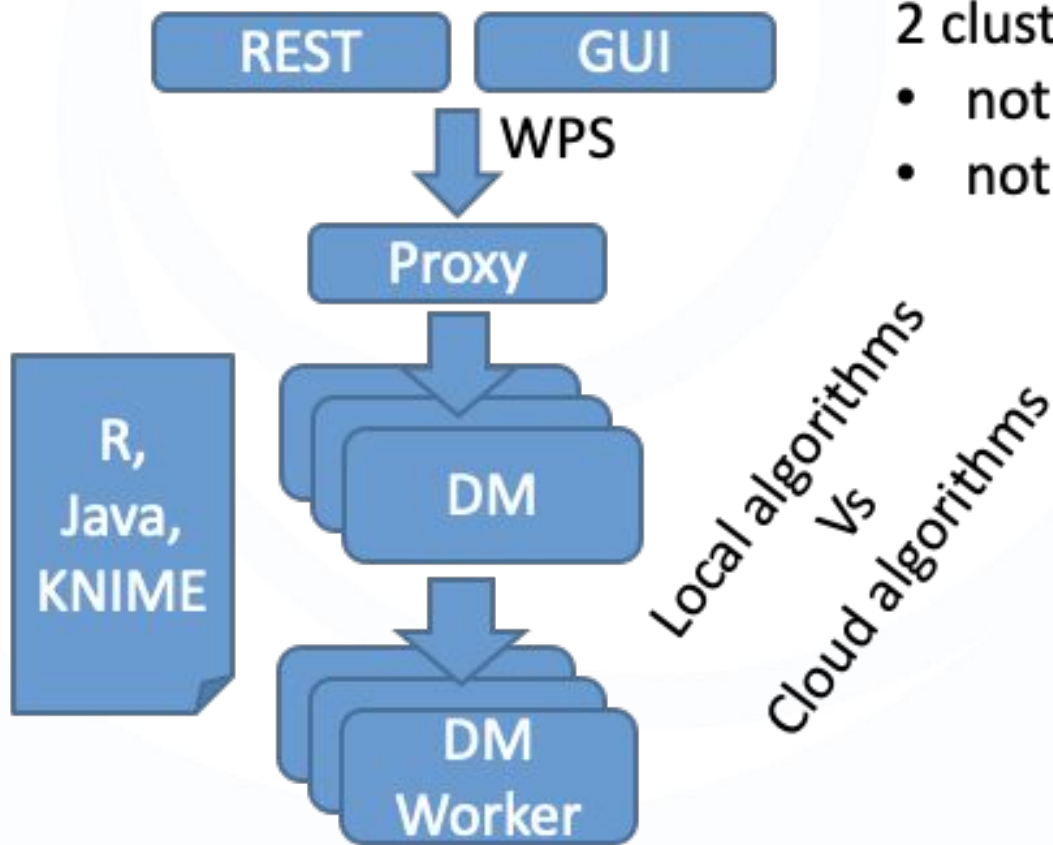
News feed: Shows posts sorted by newest Post.

- Post 1:** Catalogue, April 04, 6:23 PM. Text: Taras Günther just published the item "(surveillance) data harmonization". Please find it at: <http://data.d4science.org/ctlg/AGINFRAplusDev/f9df3fbc-2829-4ed2-a849-0de4c576639e>. Hashtags: #Animal_Health #Data_processing_and_analysis #Epidemiology #Food_Safety #One_Health_EJP_Glossary #Public_Health.
- Post 2:** Catalogue, February 13, 1:08 PM. Text: Dear members, Rob Knapen just published the item "CROP_SIMULATION_FOR_FIELD_V1_ID_43a9f471_9f95_4e6f_bb75_360ebaded77d". You can find it at: http://data.d4science.org/ctlg/AGINFRAplusDev/crop_simulation_for_field_v1_id_43a9f471_9f95_4e6f_bb75_360ebaded77d. Hashtags: #2018 #maize #limbrug.
- Post 3:** Catalogue, February 12, 11:18 AM. Text: Dear members, Leonardo Candela just published the item "D8.4 AGINFRA PLUS Sustainability Roadmap_M36.pdf". You can find it at: http://data.d4science.org/ctlg/AGINFRAplusDev/d8_4_aginfra_plus_sustainability_roadmap_m36_pdf. Hashtags: #Zenodo #Sample #abruzzo.

About: Features the AGINFRA DEV logo and text: "This Virtual Research Environment is conceived to provide AGINFRAplus consortium members with a working environment to experience with AGINFRAplus services and develop new ones." Includes an "Edit this text" button and "Other options ...".

VRE Managers and Groups: Includes a "View Managers" button.

Invite Members: Includes an "e-mail address" input field and a "Send Invite" button.

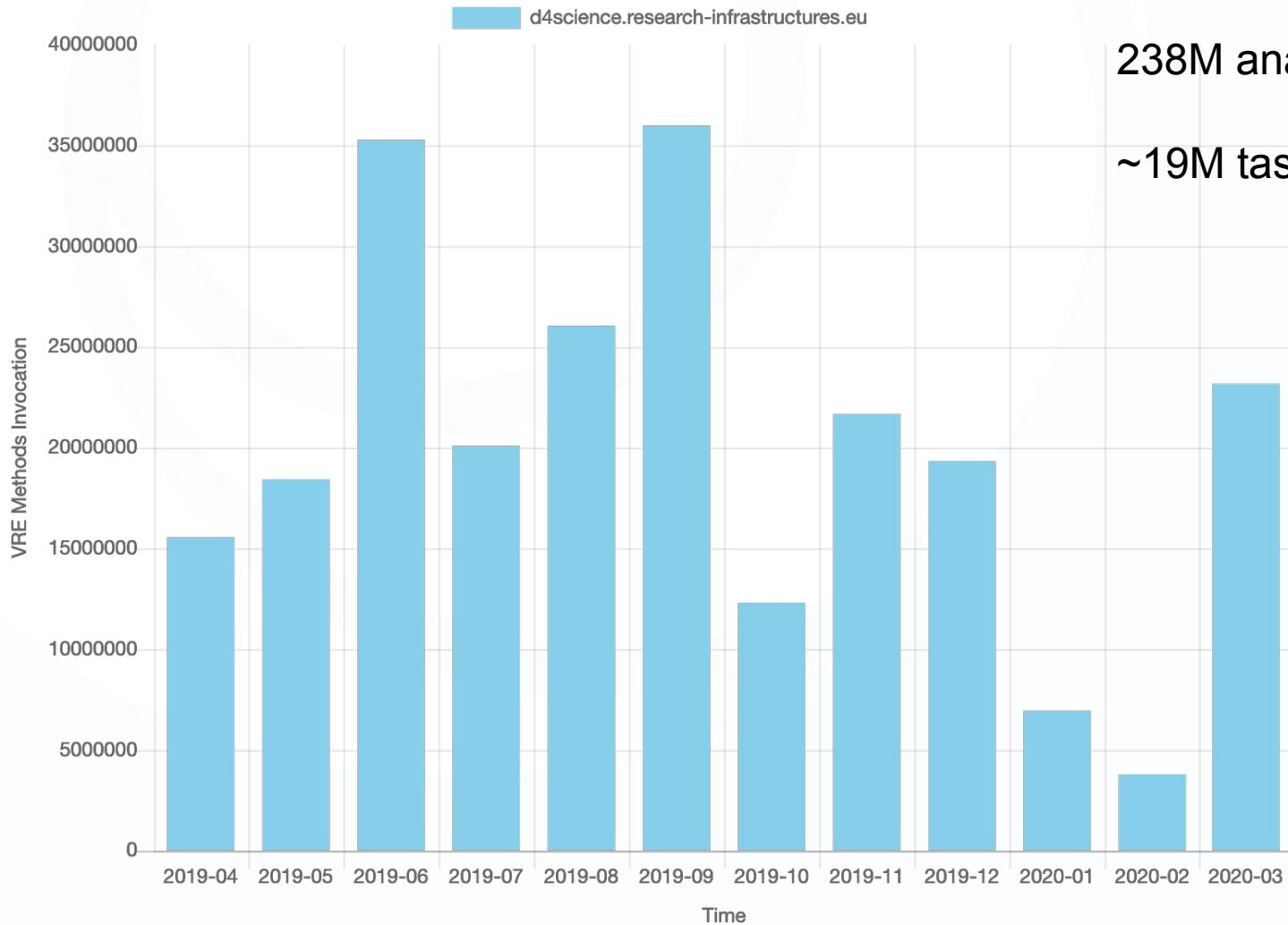


2 clusters:

- not dedicated to AGINFRA+
- not co-existing in the same VRE

Proto	Prod
6 servers <ul style="list-style-type: none"> • 16 vCores • 16 GB RAM • 100 GB Space 	15 servers <ul style="list-style-type: none"> • 16 vCores • 32 GB RAM • 100 GB Space
6 servers <ul style="list-style-type: none"> • 16 vCores • 16 GB RAM • 100 GB Space 	15 servers <ul style="list-style-type: none"> • 16 vCores • 32 GB RAM • 100 GB Space

d4science.research-infrastructures.eu VRE Methods Invocation [Method Invocation]



238M analytics tasks

~19M tasks x month



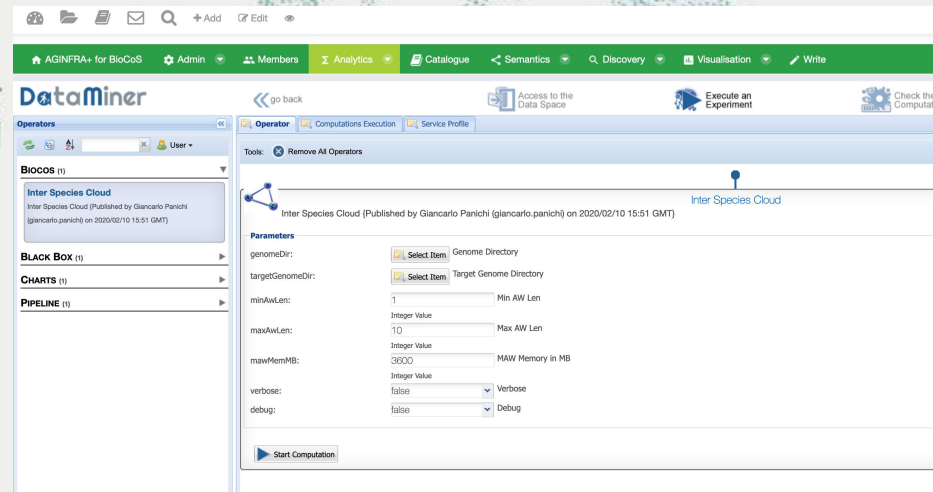
The tool allows a fast identification of species-specific genomic loci containing potential genetic markers.

The functional principle is based on the identification of sequences non-tolerated in a predefined set of genomes (Query) but evolutionary conserved in the genome of interest (Target).

It can be applied for the identification of genetic markers for inter- and intra-species analyses – in base of the biological need.

Advantages:

- 1) Overcomes whole genome alignments
- 2) Reduces sequencing needs



The screenshot shows the Dataminer web interface. The top navigation bar includes links for AGINFRA+ for BioCoS, Admin, Members, Analytics, Catalogue, Semantics, Discovery, Visualisation, and Write. The main interface displays the 'Operators' section with a search bar and a list of operators. The 'Inter Species Cloud' operator is selected, showing its parameters:

- genomeDir: (with a 'Select Item' button)
- targetGenomeDir: (with a 'Select Item' button)
- minAwLen: (with a 'Min AW Len' label and 'Integer Value' dropdown)
- maxAwLen: (with a 'Max AW Len' label and 'Integer Value' dropdown)
- maxMemMB: (with a 'MAW Memory in MB' label and 'Integer Value' dropdown)
- verbose: (with a 'Verbose' label and dropdown)
- debug: (with a 'Debug' label and dropdown)

A 'Start Computation' button is located at the bottom of the configuration panel.